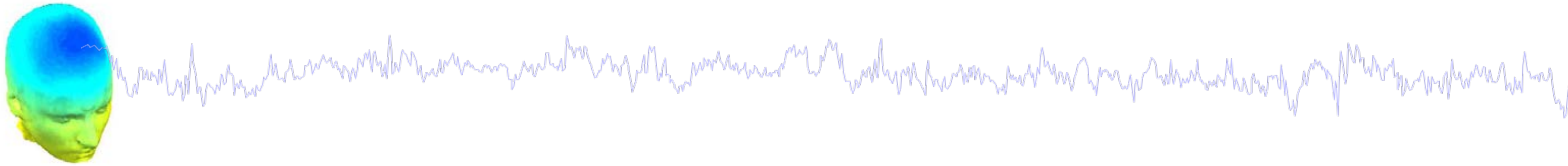


Source information flow and Granger-Causal modeling tools

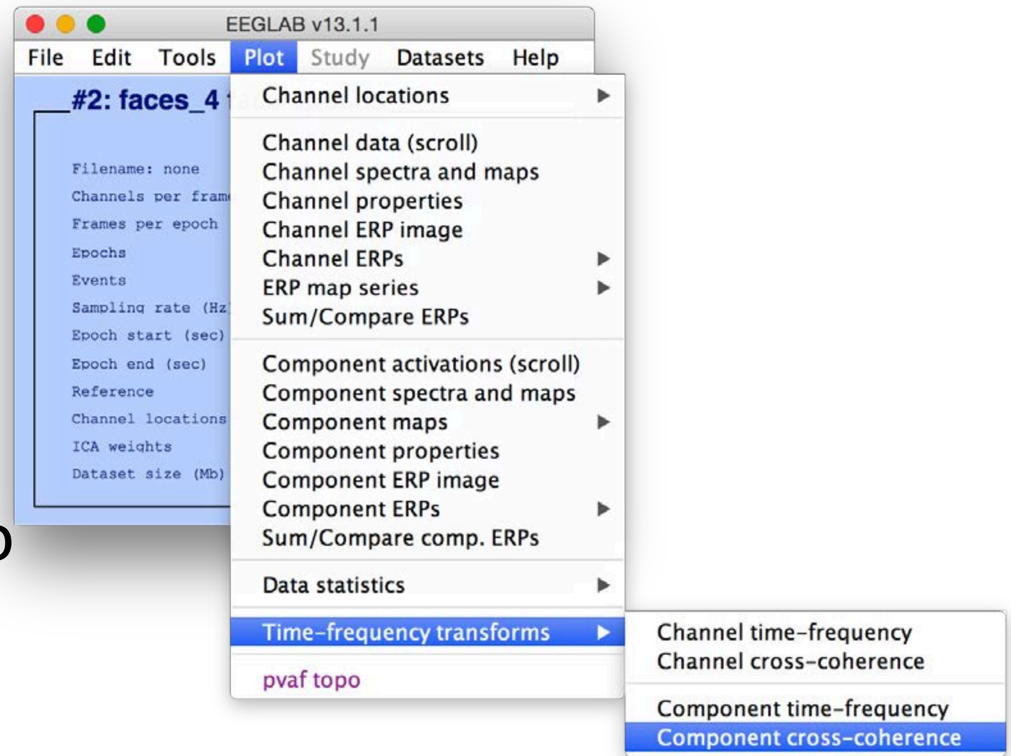
EEGLAB Workshop XXIII

AIISH, Mysuru, India

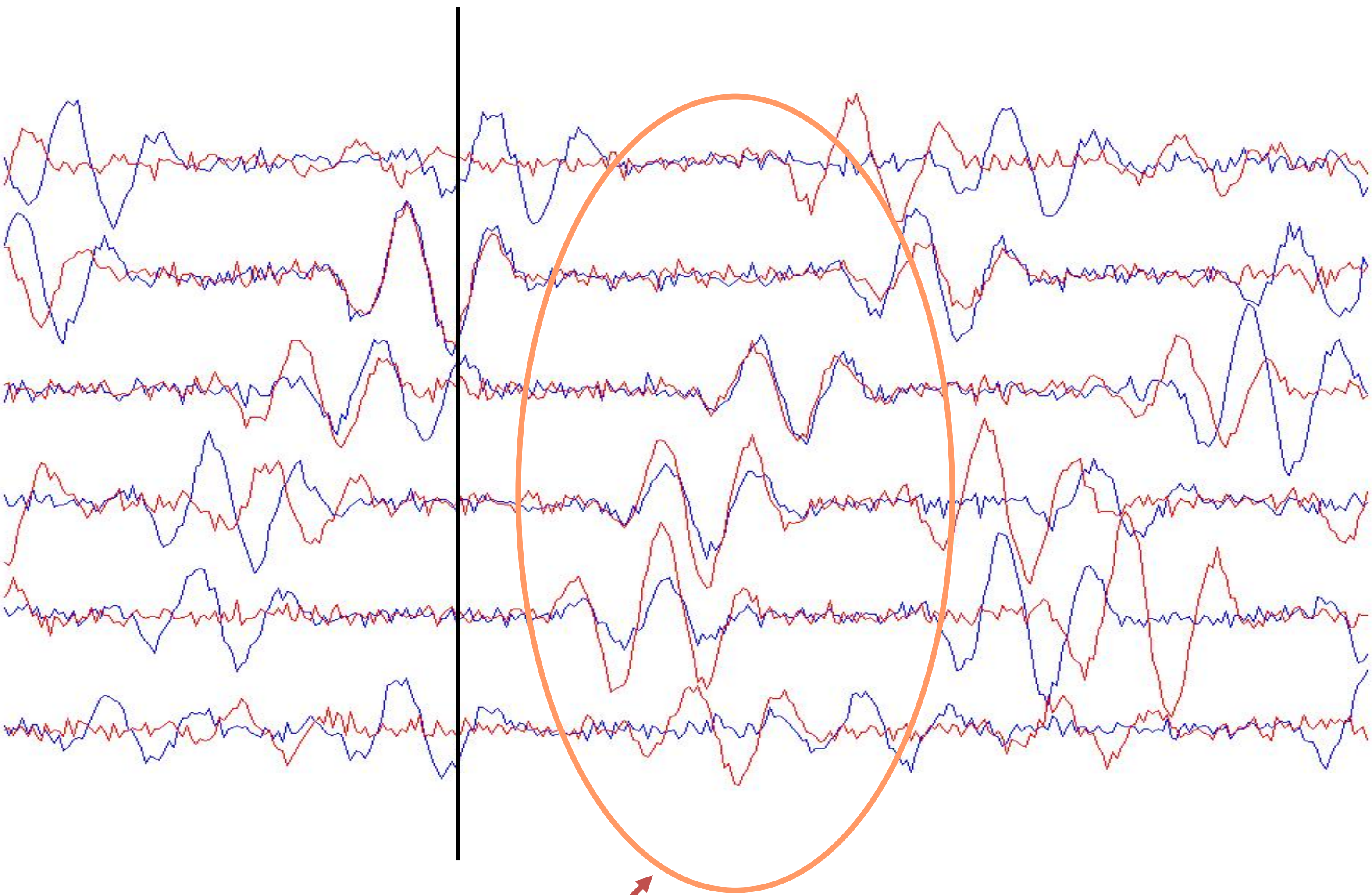
Day 3

Part 3b: Event Related Coherence

- Goal: How similar is the event-related response of two signals?
  - Between channels (problematic due to volume conduction)
  - Between ICs
  - Useful to quickly begin to understand relationships between components

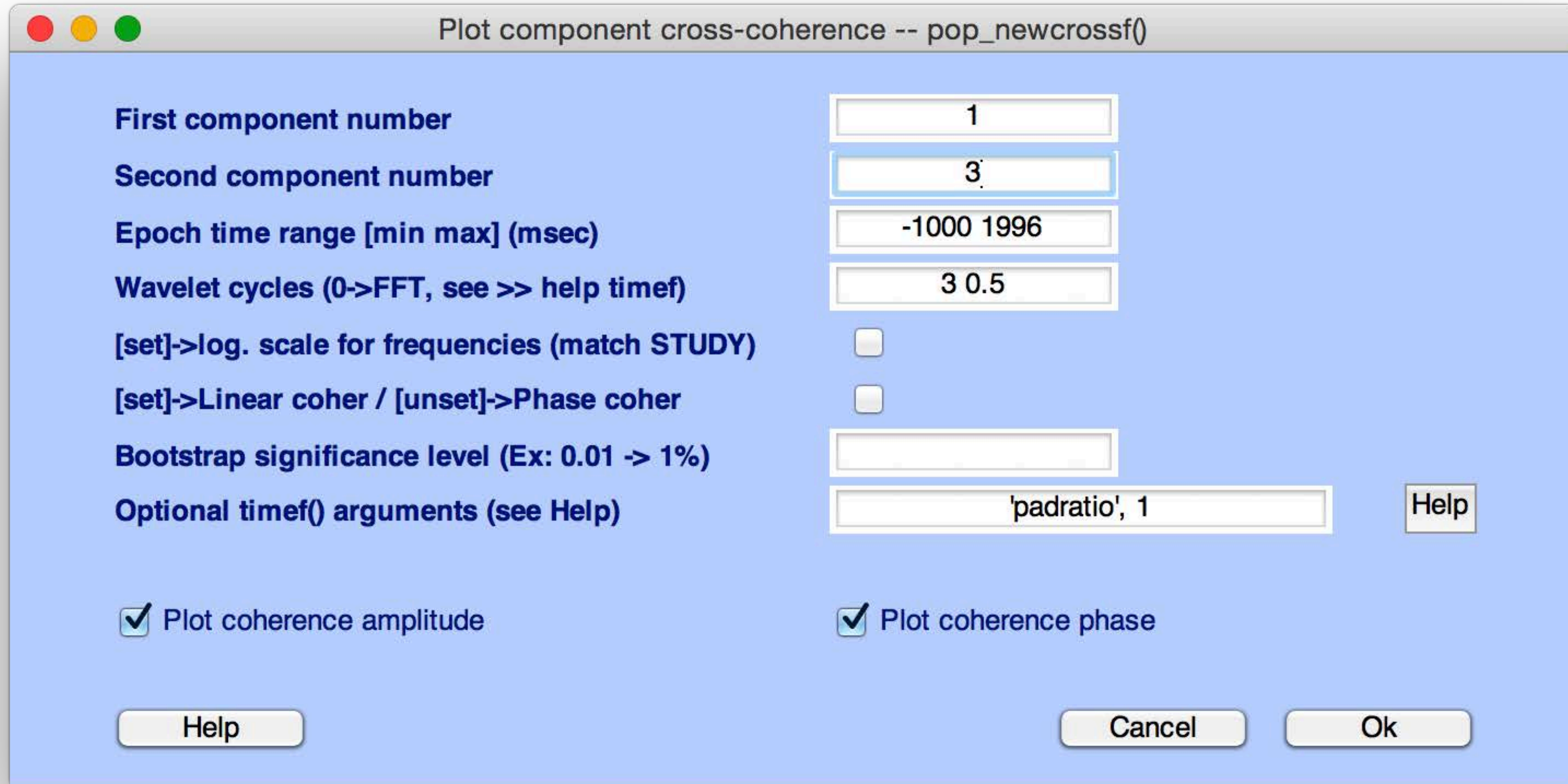


TWO SIMULATED THETA PROCESSES



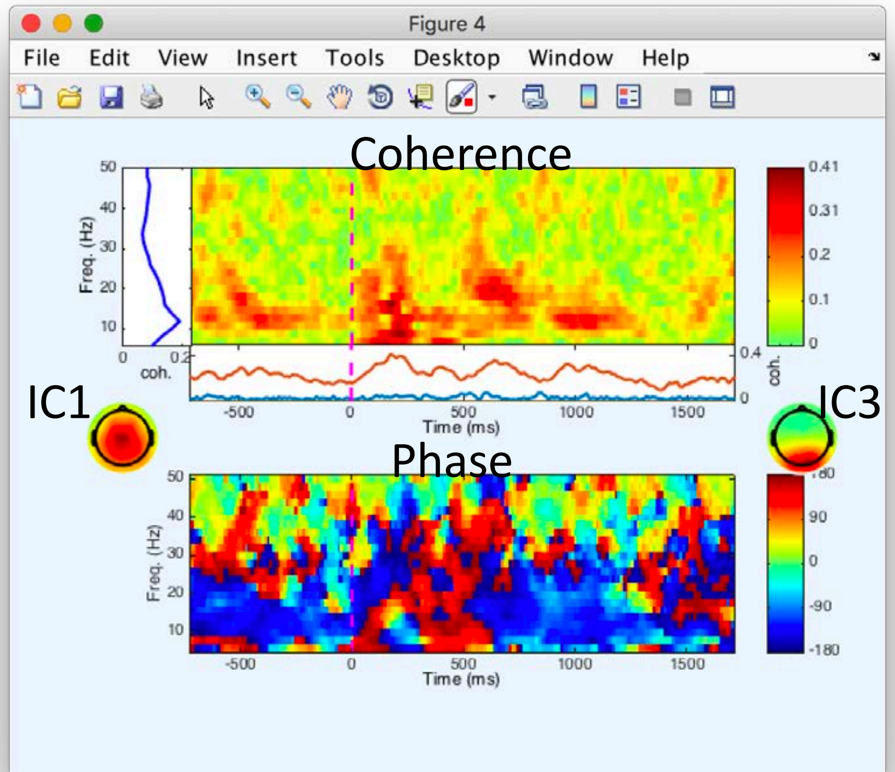
Event-related Coherence

Try it!

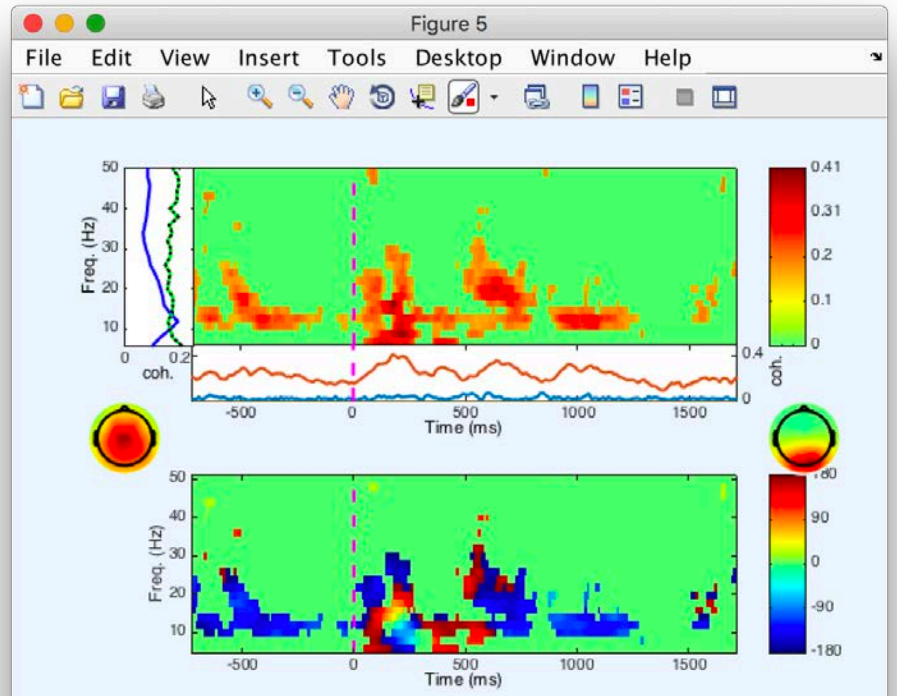

Plot component cross-coherence -- pop_newcrossf()
First component number
1
Second component number
3
Epoch time range [min max] (msec)
-1000 1996
Wavelet cycles (0->FFT, see >> help timef)
3 0.5
[set]->log. scale for frequencies (match STUDY)
[set]->Linear coher / [unset]->Phase coher
Bootstrap significance level (Ex: 0.01 -> 1%)
Optional timef() arguments (see Help)
'padratio', 1
Help
Plot coherence amplitude
Plot coherence phase
Help
Cancel
Ok

4

Cross coherence between IC 1 and IC 3



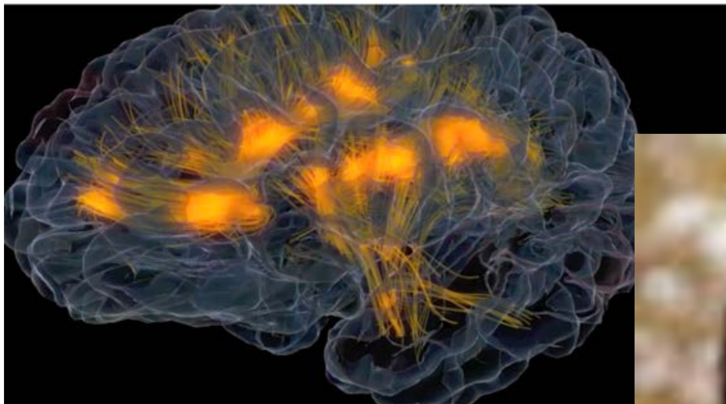
$\alpha = 0.01$



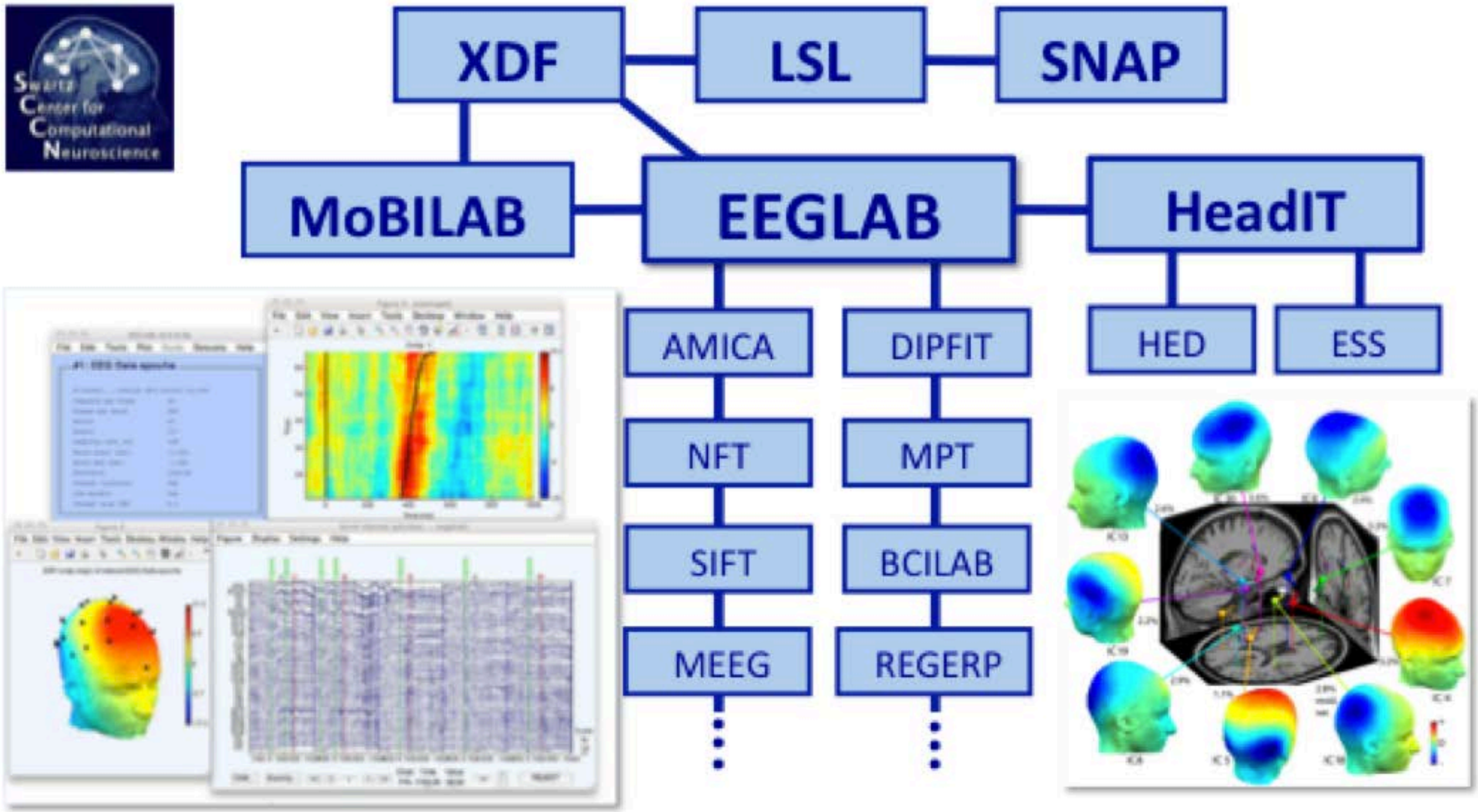
Significant event-related coherence (as well as tonic coherence) in alpha/beta bands
IC 1 tonically leads IC 3 (negative phase), but phase relationships are changed post-stimulus

Directional measures of effective connectivity are present in the SIFT toolbox.

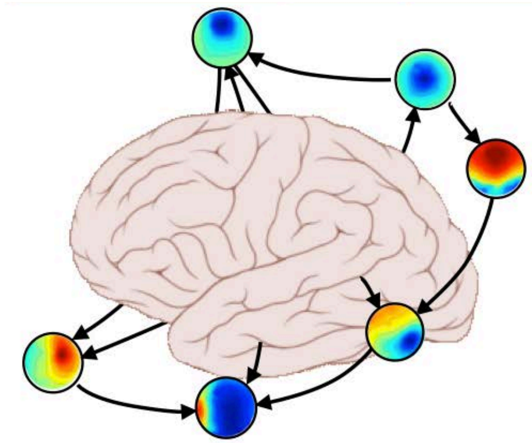
Tim Mullen



EEGLAB Toolset



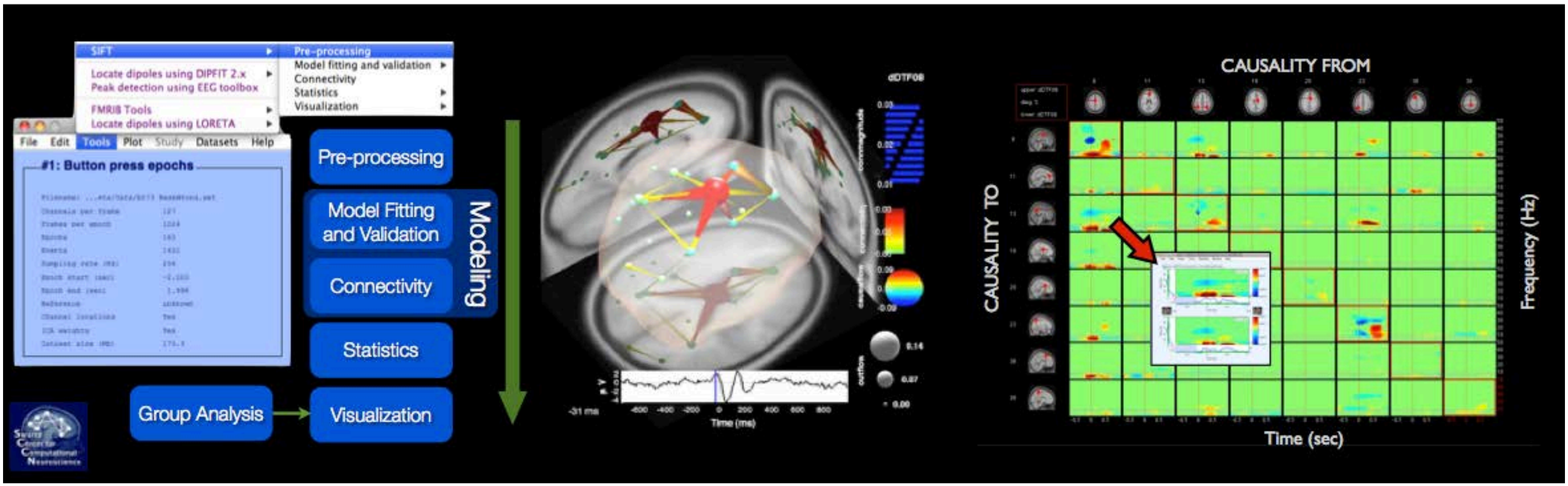
http://sccn.ucsd.edu/eeglab/



SIFT

Source Information Flow Toolbox

http://sccn.ucsd.edu/wiki/SIFT
Mullen, et al, *Journal of Neuroscience Methods* (in prep, 2012)
Mullen, et al, *Society for Neuroscience*, 2010
Delorme, Mullen, Kothe et al, *Computational Intelligence and Neuroscience*, vol 12, 2011



- A toolbox for (source-space) electrophysiological information flow and causality analysis (single- or multi-subject) integrated into the EEGLAB software environment.
- Emphasis on vector autoregression and time-frequency domain approaches
- Standard and novel interactive visualization methods for exploratory analysis of connectivity across time, frequency, and spatial location

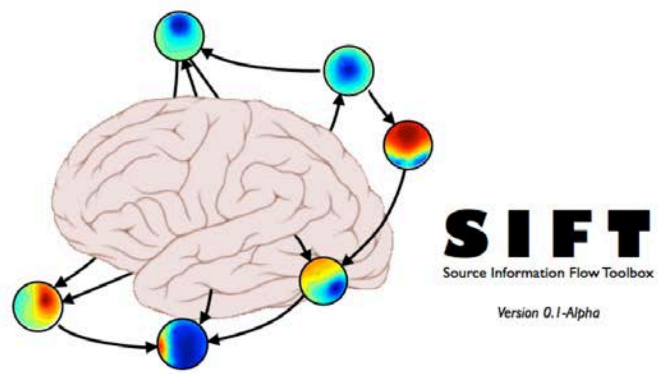
sccn.ucsd.edu/wiki/SIFT

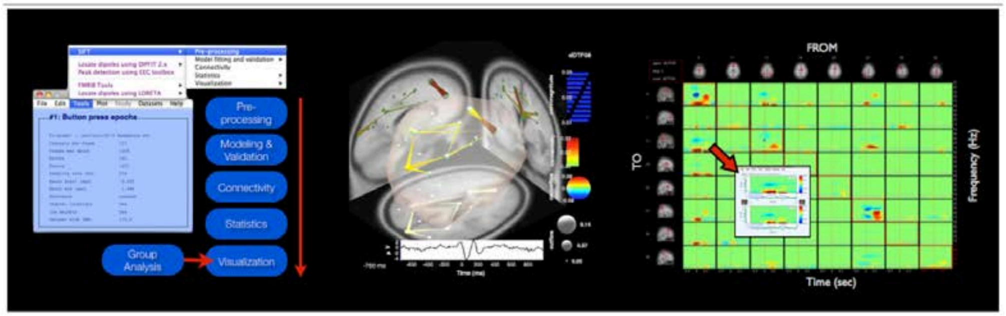
Zoom on ucsd nsi SCCN Scholar Wikipedia EN Popular OCR whatstat Nolte EEGLAB Undoc Matlab News PowerSchool neo lit

Log in

page discussion view source history

SIFT





Contents [hide]

Welcome to the repository for the Source Information Flow Toolbox (SIFT)

Developed and Maintained by: Tim Mullen (SCCN, INC, UCSD)
Web: http://www.antillipsi.net
Email: <Tim's first name> (at) sccn (dot) ucsd (dot) edu

home
- SCCN web site
- EEGLAB Wiki
- MoBI Lab Wiki
- SCCN Wiki Home

eeglab pages
- EEGLAB Home
- EEGLAB Wiki
- EEGLAB Tutorial
- Online EEGLAB Workshop
- Download EEGLAB
- Revision history
- Help EEGLAB

sccn toolboxes
- EEGLAB
- NFT
- BCILAB
- SIFT
- MoBILAB
- MPT

wiki tools
- Sandbox
- Basic Wiki Syntax
- Wiki Help
- New Users
- Recent changes

search

Search | Go | Search

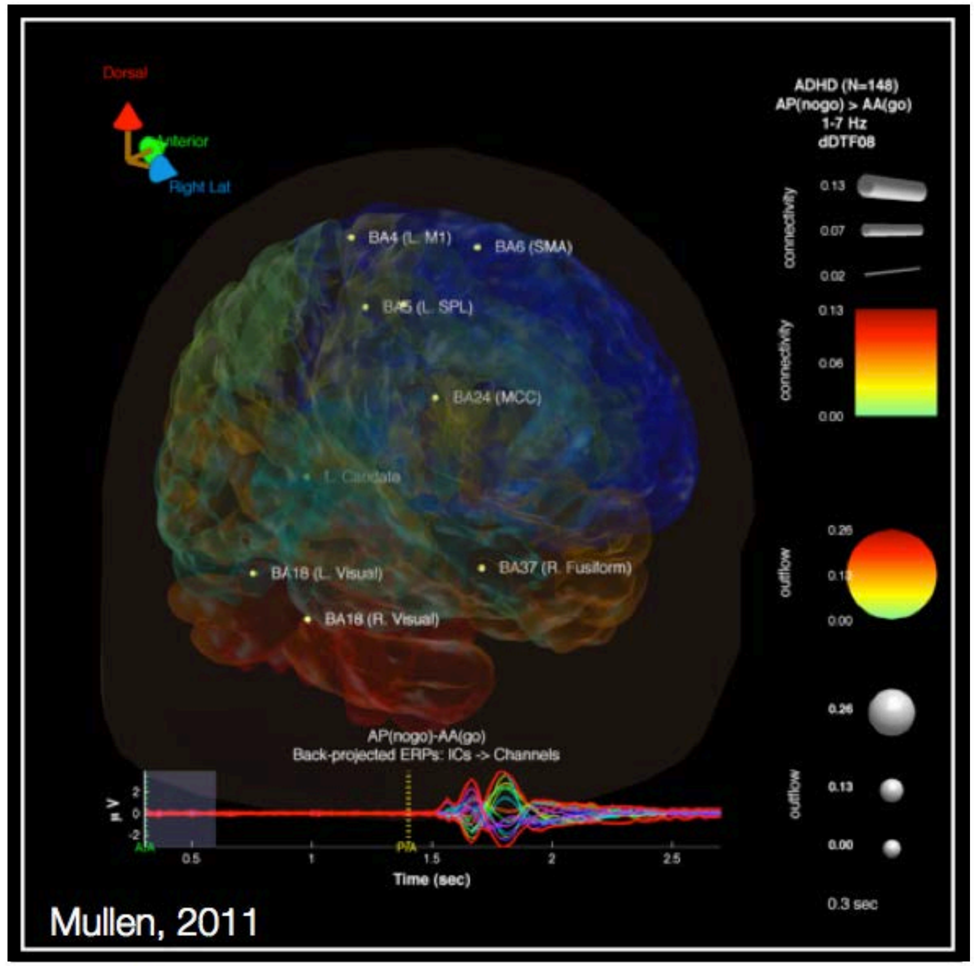
advanced wiki tools
- What links here
- Related changes
- Special pages
- Printable version
- Permanent link
- Page information

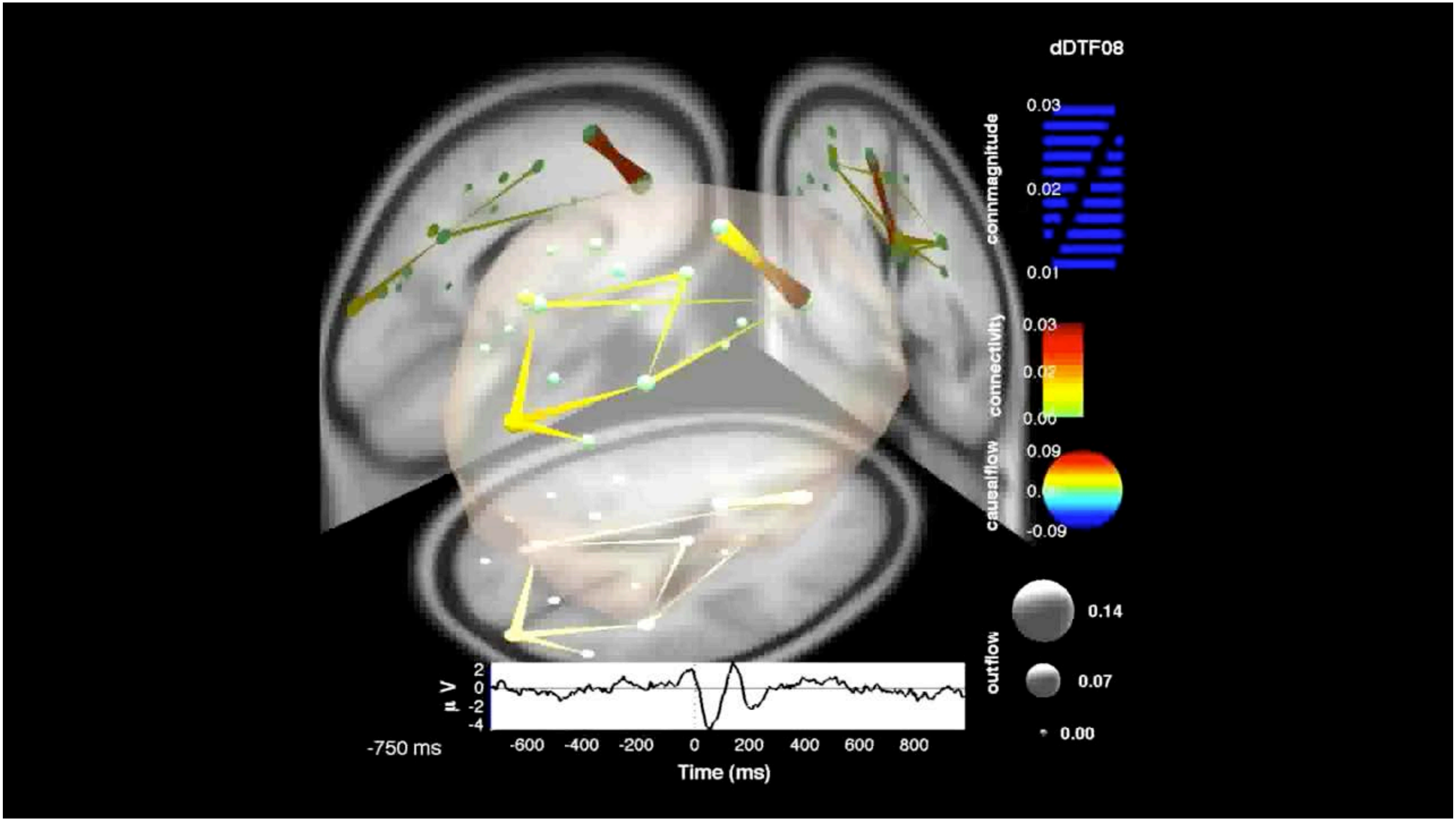
Display a menu

The Dynamic Brain

- A key goal: To model temporal changes in neural dynamics and information flow that index and predict task-relevant changes in cognitive state and behavior

- **Open Challenges:**
  - Non-invasive measures (source inference)
  - Robustness and Validity (constraints & statistics)
  - Scalability (multivariate)
  - Temporal Specificity / Non-stationarity / Single-trial (dynamics)
  - Multi-subject Inference
  - Usability and Data Visualization (software)

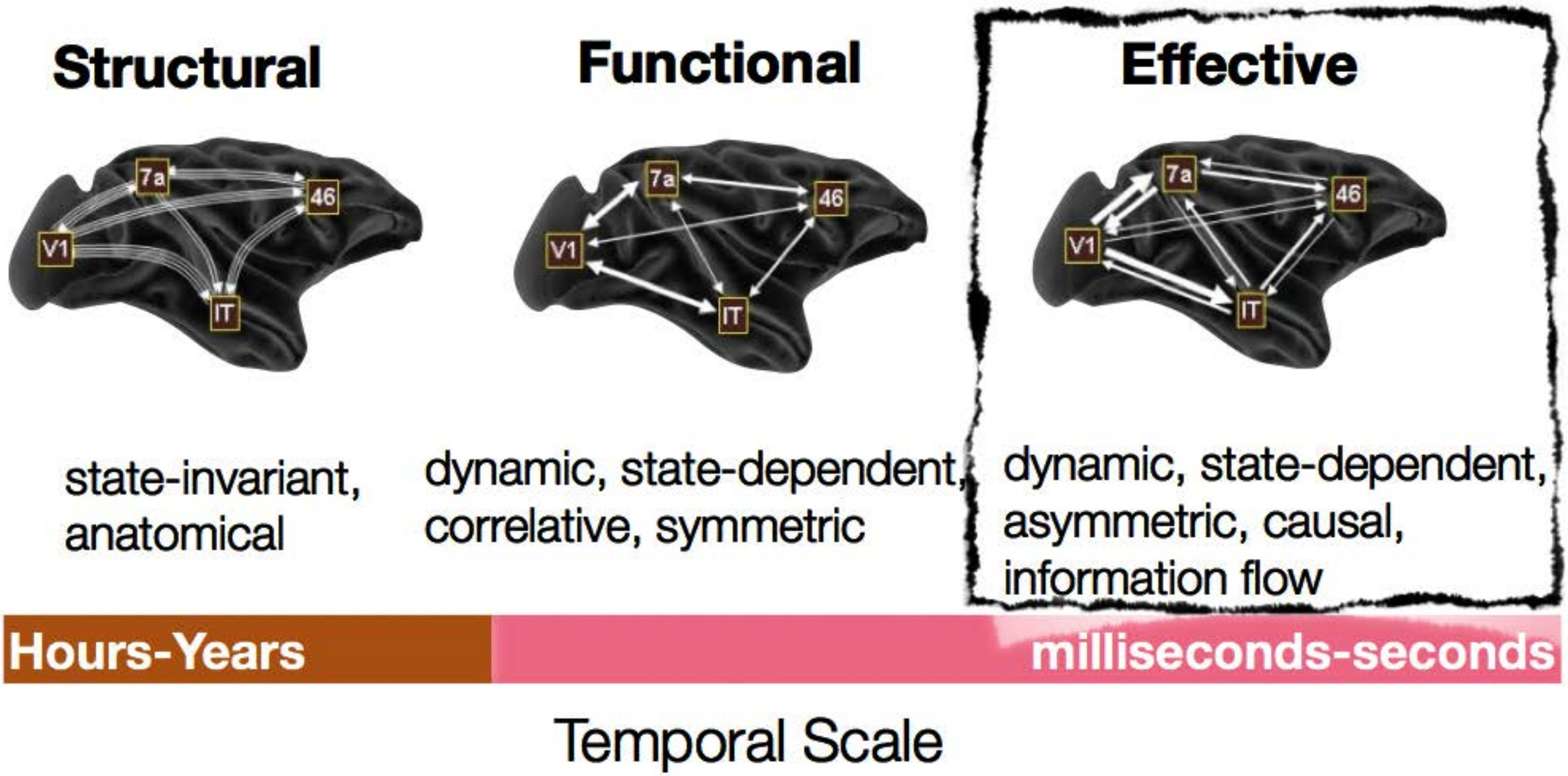


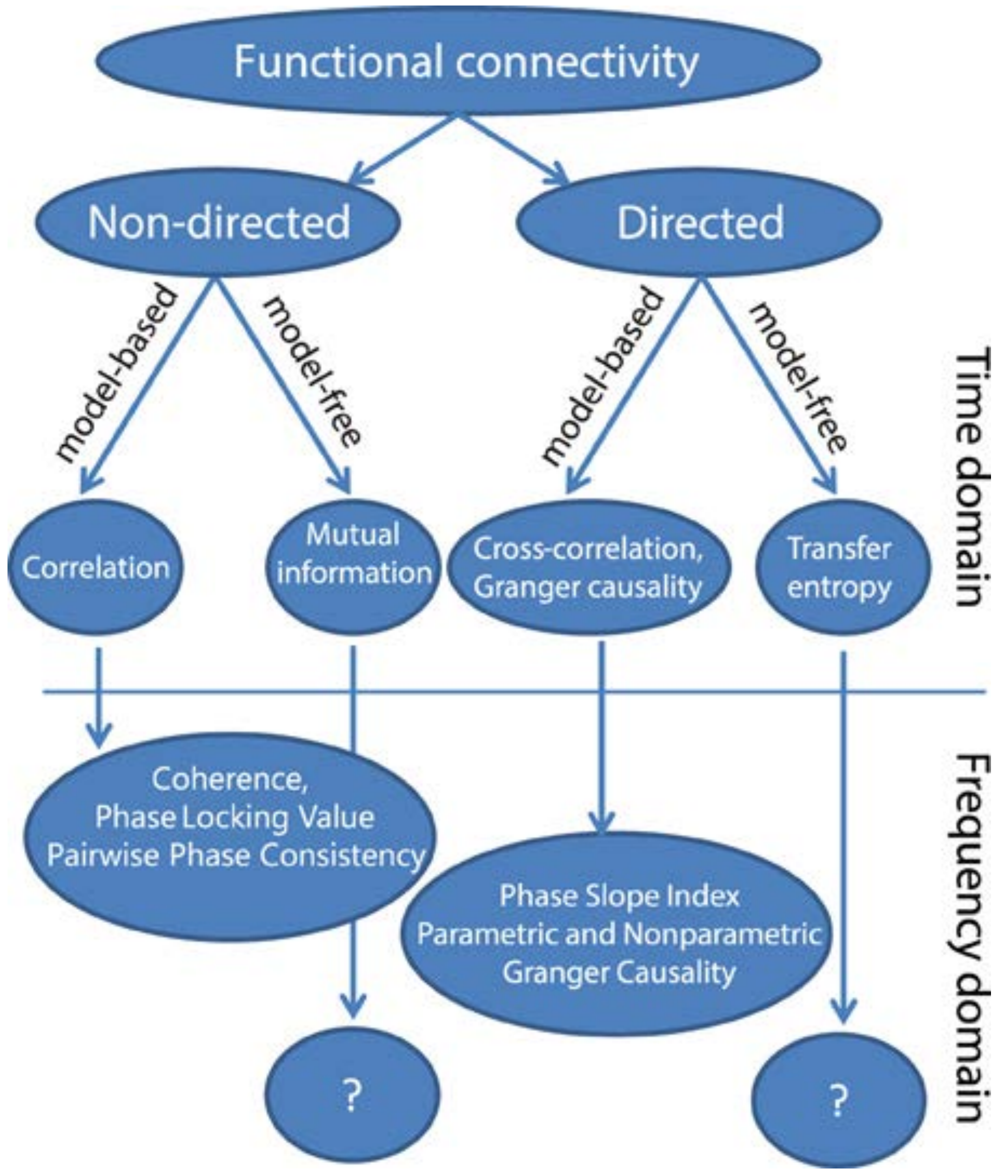

dDTF08
connmagnitude
0.03
0.02
0.01
connectivity
0.03
0.00
causalflow
0.09
-0.09
outflow
0.14
0.07
0.00
μV
2
0
-2
-4
-750 ms
-600
-400
-200
0
200
400
600
800
Time (ms)

Tim Mullen

Large-scale brain connectivity

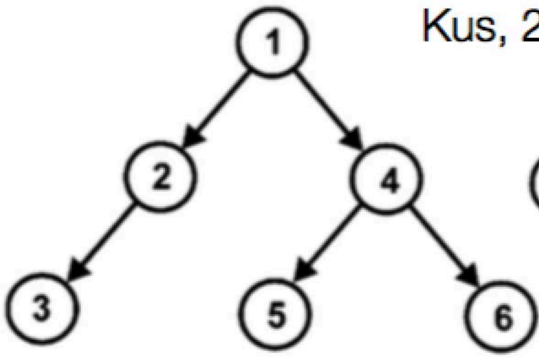
(Bullmore and Sporns, *Nature*, 2009)





Bastos AM, Schoffelen J-M: **A Tutorial Review of Functional Connectivity Analysis Methods and Their Interpretational Pitfalls**. *Front Sys Neurosci* 2016, **9**:413.

The problem of spurious connectivity

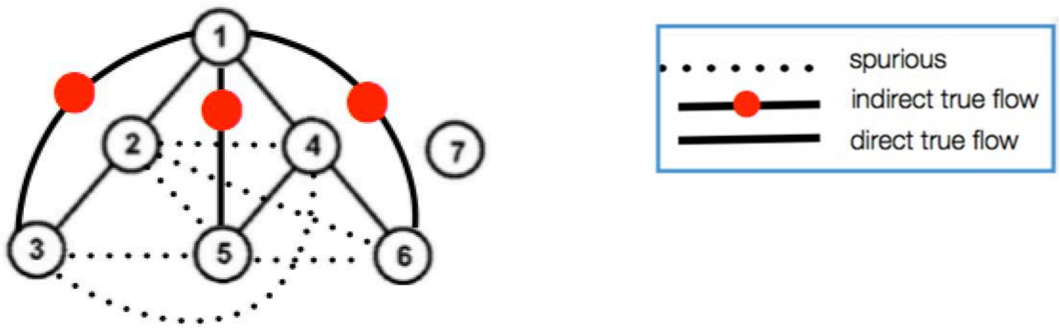


Kus, 2004

Coherency

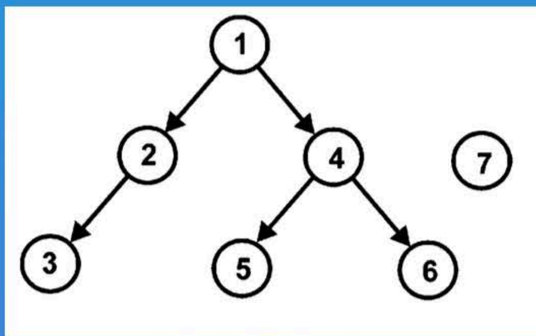
$$C_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f)S_{jj}(f)}}$$

(Bendat and Piersol, 1986)

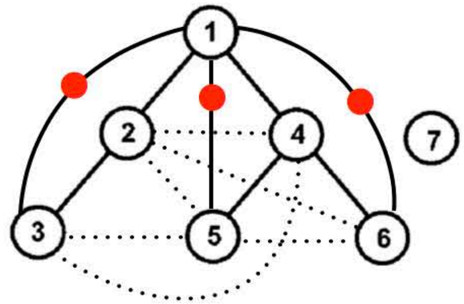


Bivariate measures, such as coherence (but also original GC), find spurious connections between nodes if they share a common input.

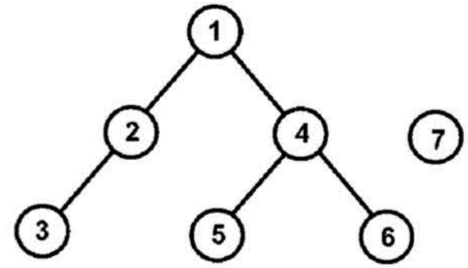
Ground Truth



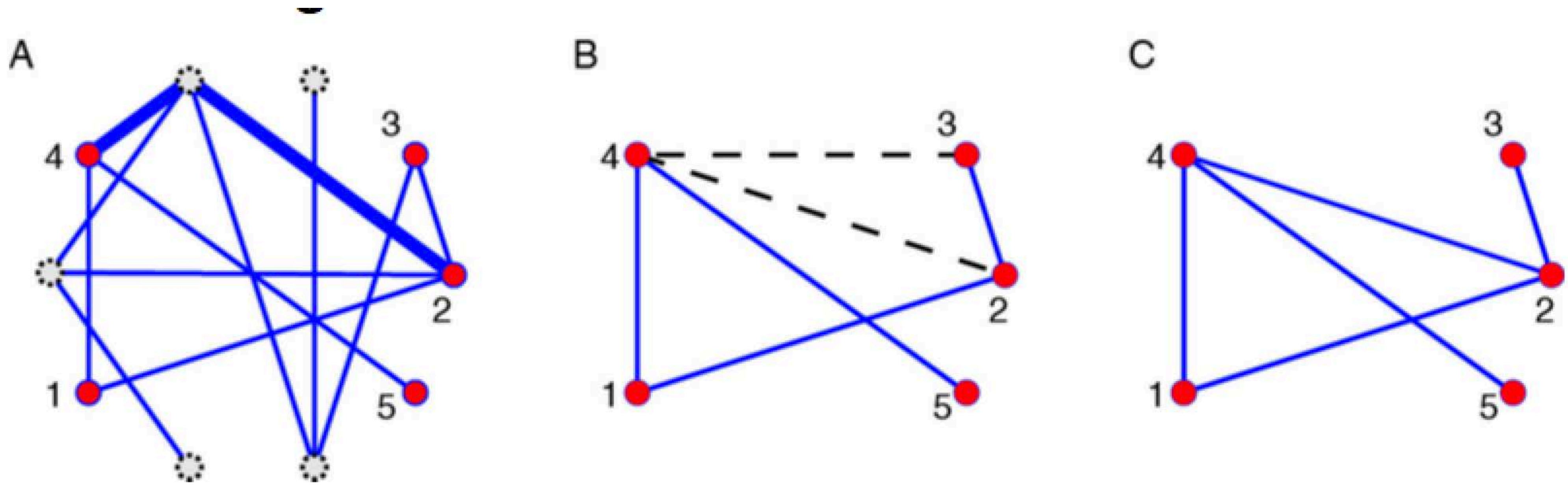
Coherence



Partial Coherence



A deeper problem – unobserved nodes



With EEG, it's unavoidable that there will be contributing network nodes (e.g. thalamus) that we cannot observe.

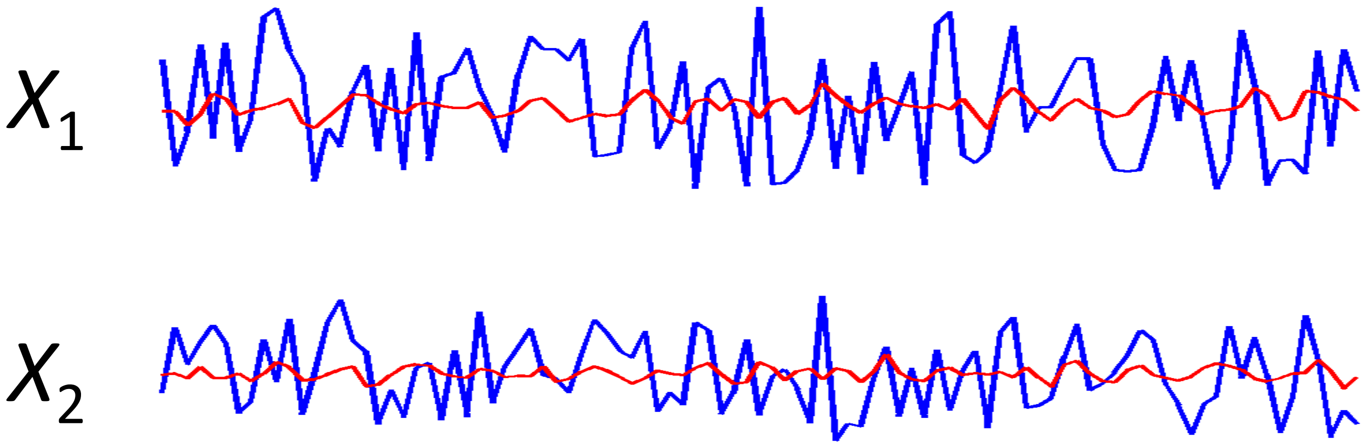
We also can't be sure ICA will identify all important sources…

Granger-causality

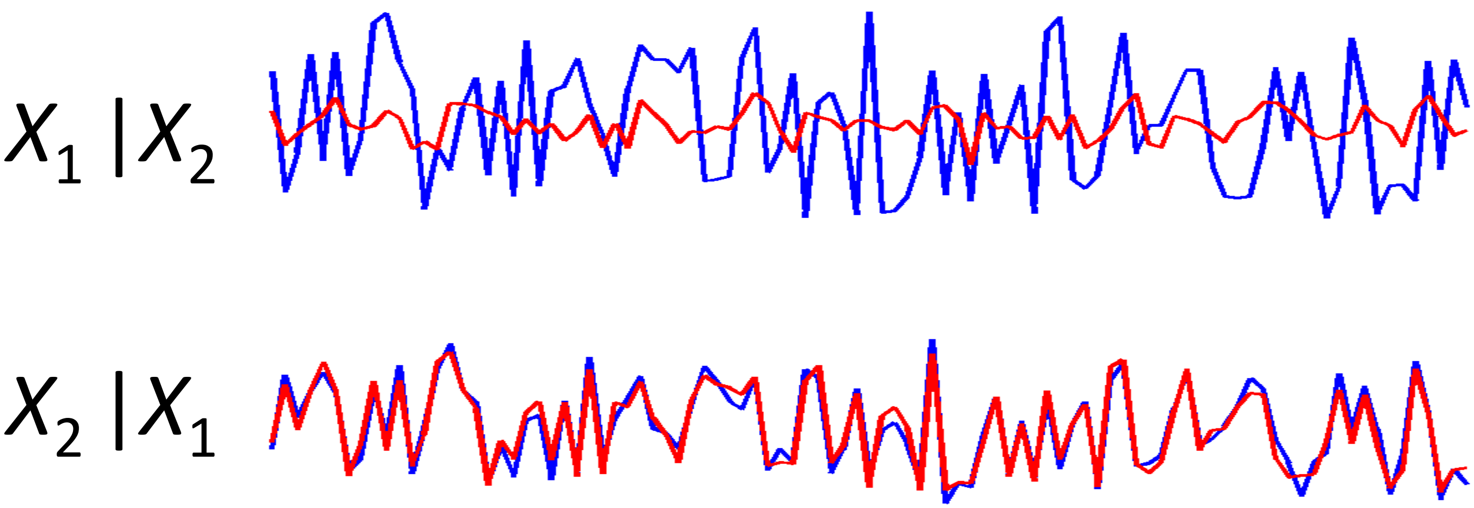
- A measure of *statistical* causality based on prediction.
- Widely used in time-series econometrics.
- Nobel Prize in economics, 2003.

If a signal A causes a signal B, then knowledge of the past of both A and B should improve the predictability of B, as compared to knowledge of B alone.

AR Models (prediction of future of a signal by its past)



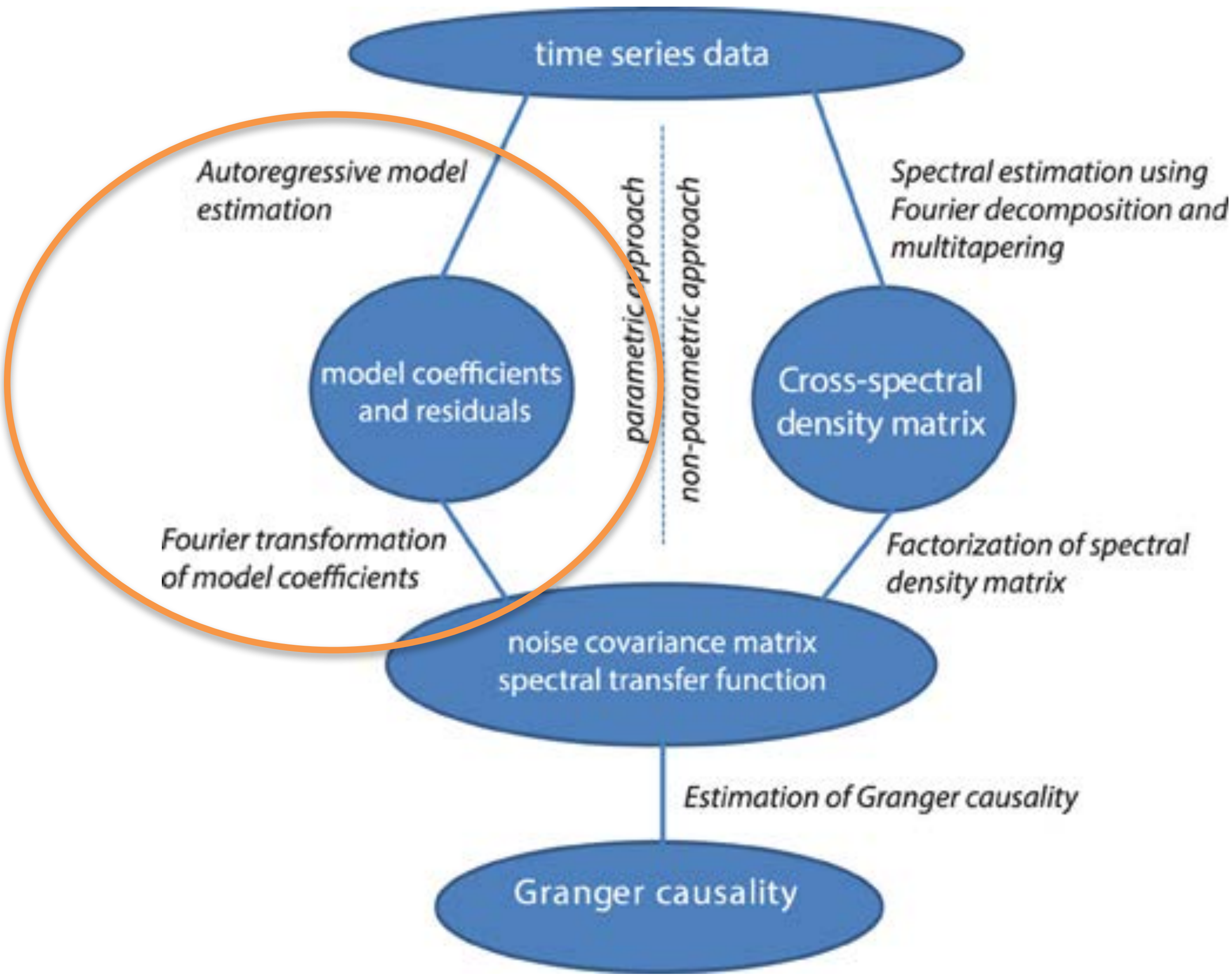
VAR Models (prediction of future of a signal by its past + the other signal's past)



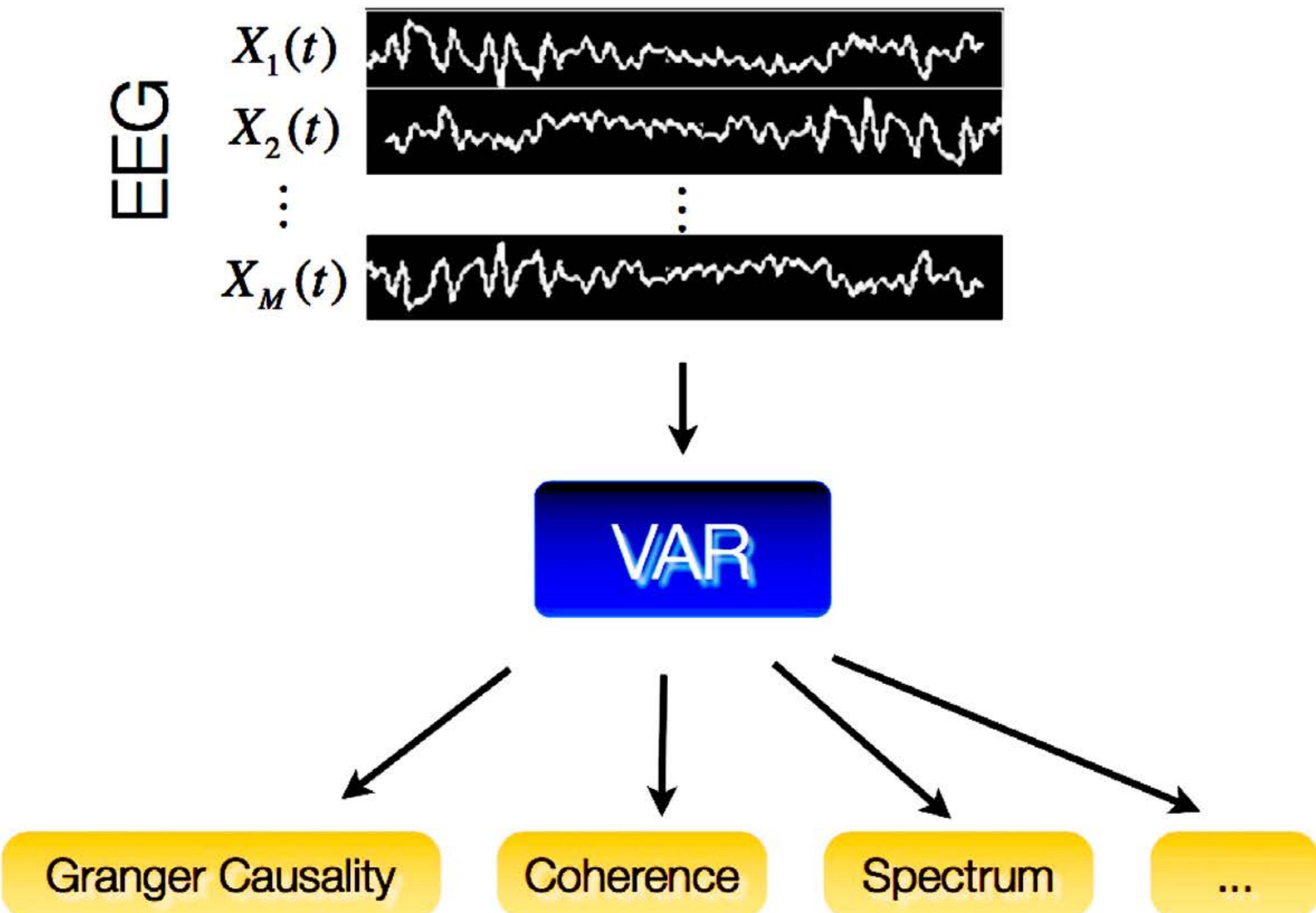
Incorporating information about X_1 improves the prediction of X_2!

We say "X_1 granger-causes X_2"

Calculation of GC

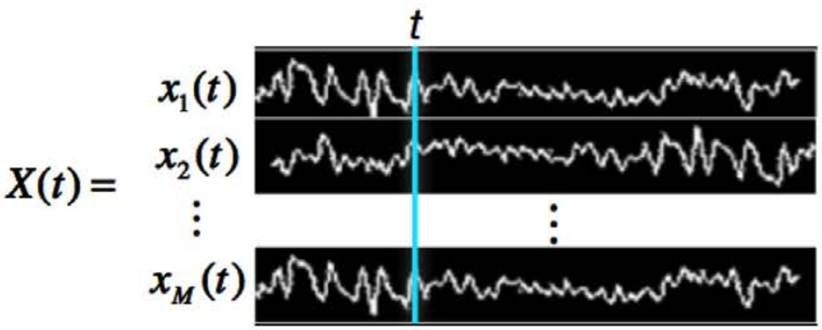


Vector Autoregressive (VAR / MAR / MVAR) Modeling



The Linear Vector Auto-regressive (VAR) Model

Ordinary Least-Squares



VAR[p] model

$$\mathbf{X}(t) = \sum_{k=1}^{p} \mathbf{A}^{(k)}(t)\mathbf{X}(t-k) + \mathbf{E}(t)$$

model order

random noise process

M-channel data vector at current time t

M x M matrix of (possibly time-varying) model coefficients indicating variable dependencies at lag k

multichannel data k samples in the past

$$\mathbf{A}^{(k)}(t) = \begin{pmatrix} a_{11}^{(k)}(t) & \cdots & a_{1M}^{(k)}(t) \\ \vdots & \ddots & \vdots \\ a_{M1}^{(k)}(t) & \cdots & a_{MM}^{(k)}(t) \end{pmatrix} \qquad \mathbf{E}(t) = N(0, \mathbf{V})$$

Selecting a VAR Model Order

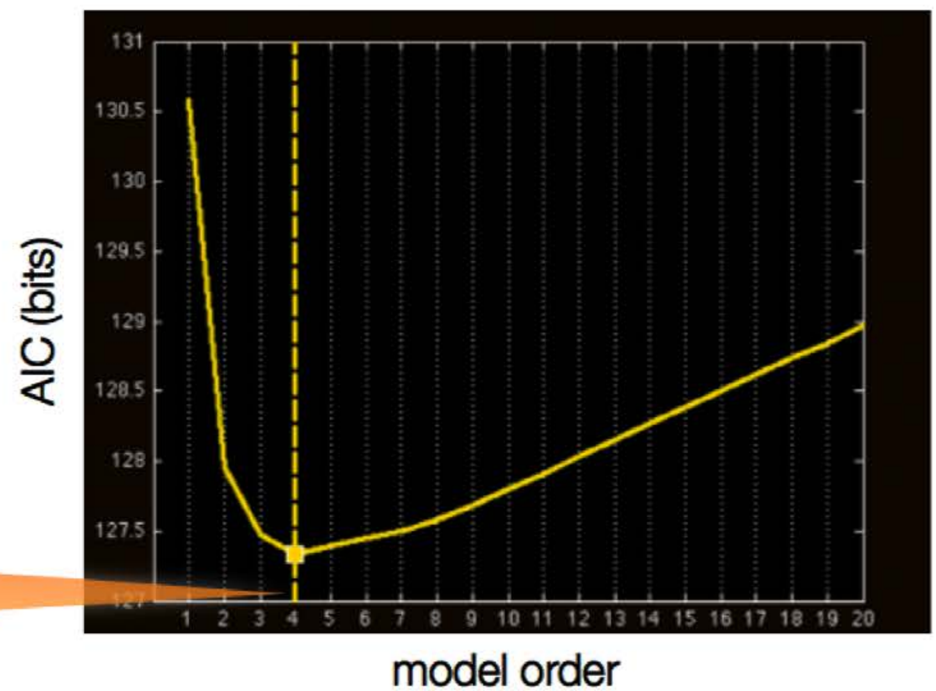
- Model order is typically determined by minimizing information criteria such as Akaike Information Criterion (AIC) for varying model order (p):

$$AIC(p) = 2\log(\det(\mathbf{V})) + M^2p/N$$

Penalizes high model orders (parsimony)

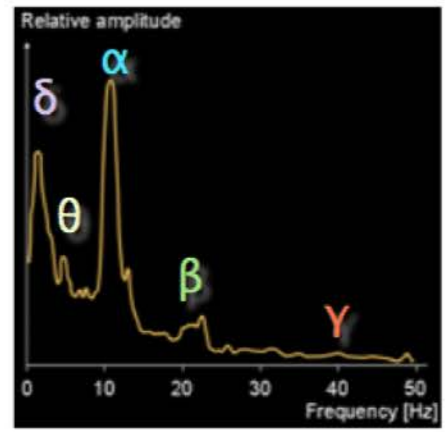
entropy rate (amount of prediction error)

optimal order



Selecting a VAR Model Order

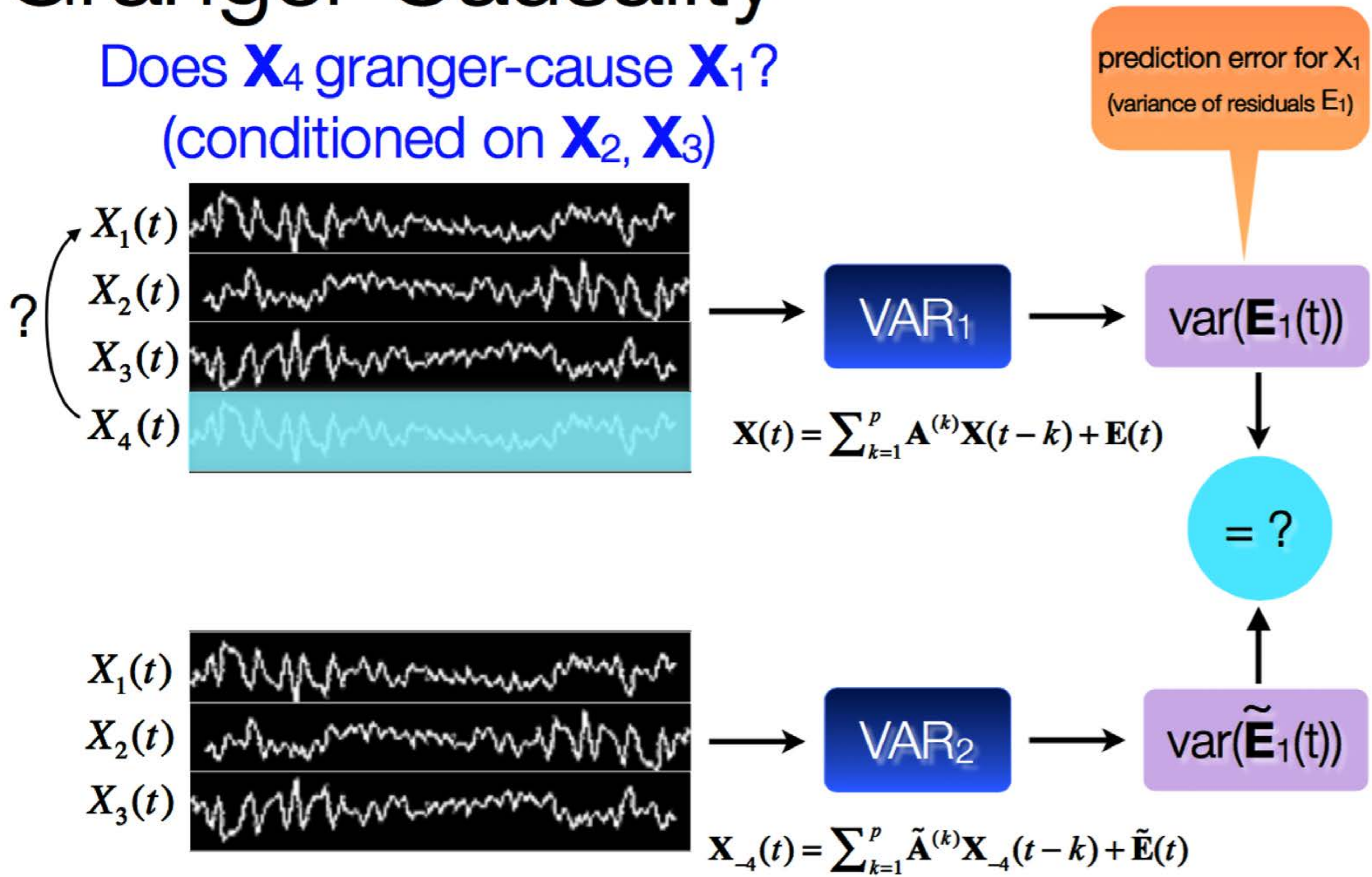
- Other considerations:
  - A M-dimensional VAR model of order *p* has at most *Mp/2* spectral peaks distributed amongst the *M* variables. This means we can observe at most *p/2* peaks in each variables' spectrum (or in the causal spectrum between each pair of variables)



  - Optimal model order depends on sampling rate (higher sampling rate often requires higher model orders)

Granger Causality

Does $\mathbf{X}_4$ granger-cause $\mathbf{X}_1$?
(conditioned on $\mathbf{X}_2$, $\mathbf{X}_3$)



Granger Causality

- Granger (1969) quantified this definition for **bivariate** processes in the form of an F-ratio:

$$F_{X_1 \leftarrow X_2} = \ln\left(\frac{\overbrace{var(\tilde{E}_1)}^{\text{reduced model}}}{\underbrace{var(E_1)}_{\text{full model}}}\right) = \ln\left(\frac{var(X_1(t)\,|\,X_1(\cdot))}{var(X_1(t)\,|\,X_1(\cdot), X_2(\cdot))}\right)$$

- Alternately, for a **multivariate interpretation** we can fit a single MVAR model to all channels and apply the following definition:

Definition 1

X_j granger-causes X_i conditioned on all other variables in $\mathbf{X}$

if and only if $\mathbf{A}_{ij}(k) >> 0$ for some lag $k \in \{1, \ldots, p\}$

Granger Causality – Frequency Domain

$$\mathbf{X}(t) = \sum_{k=1}^{p} \mathbf{A}^{(k)}\mathbf{X}(t-k) + \mathbf{E}(t)$$

Fourier-transforming $\mathbf{A}^{(k)}$ we obtain

Likewise, $\mathbf{X}(f)$ and $\mathbf{E}(f)$ correspond to the fourier transforms of the data and residuals, respectively

$$\mathbf{A}(f) = -\sum_{k=0}^{p} \mathbf{A}^{(k)} e^{-i2\pi fk}; \mathbf{A}^{(0)} = I$$

We can then define the spectral matrix $\mathbf{X}(f)$ as follows:

$$\mathbf{X}(f) = \mathbf{A}(f)^{-1}\mathbf{E}(f) = \mathbf{H}(f)\mathbf{E}(f)$$

Where $\mathbf{H}(f)$ is the *transfer matrix* of the system.

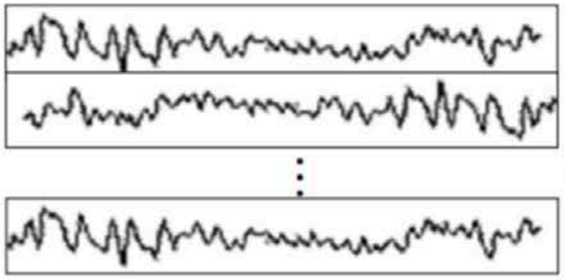
Definition 2

X_j granger-causes X_i conditioned on all other variables in $\mathbf{X}$

<u>if and only if</u> $|\mathbf{A}_{ij}(f)| >> 0$ for some frequency f

leads to PDC

$x_1(t)$
$x_2(t)$
$\vdots$
$x_M(t)$

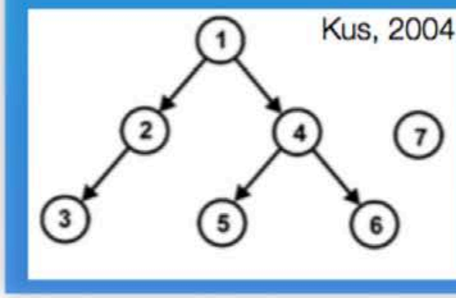


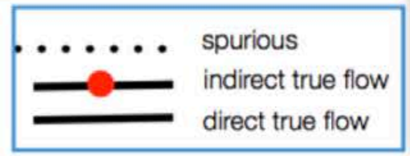
$$\mathbf{X}(t) = \sum_{k=1}^{p} \mathbf{A}^{(k)}(t)\mathbf{X}(t-k) + \mathbf{E}(t)$$

$$\mathbf{A}(f,t) = -\sum_{k=0}^{p} \mathbf{A}^{(k)}(t)e^{-i2\pi fk}; \quad \mathbf{A}^{(0)} = I$$

$$\mathbf{X}(f,t) = \mathbf{A}(f,t)^{-1}\mathbf{E}(f,t) = \mathbf{H}(f,t)\mathbf{E}(f,t)$$

Ground Truth

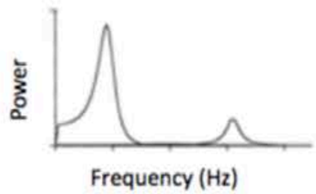




Spectrum

$$S(f) = \mathbf{X}(f)\mathbf{X}(f)^{*}$$
$$= \mathbf{H}(f)\Sigma\mathbf{H}(f)^{*}$$

(Brillinger, 2001)



NOTE: time index (t) dropped for convenience

Functional

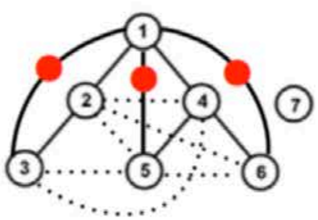
Effective

Bivariate

Coherency

$$C_{ij}(f) = \frac{S_{ij}(f)}{\sqrt{S_{ii}(f)S_{jj}(f)}}$$

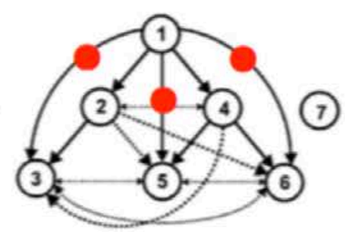
(Bendat and Piersol, 1986)



Granger-Geweke Causality

$$F_{ij}(f) = \frac{\Sigma_{jj} - (\Sigma_{ij}^2 / \Sigma_{ii}))\,|H_{ij}(f)|^2}{S_{ii}(f)}$$

(Geweke, 1982; Bressler et al., 2007)

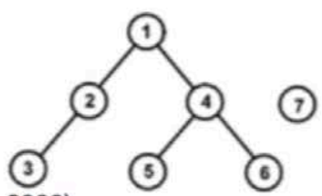


Multivariate

Partial Coherence

$$P_{ij}(f) = \frac{S_{ij}^{-1}(f)}{\sqrt{S_{ii}^{-1}(f)S_{jj}^{-1}(f)}}$$

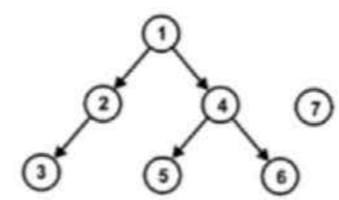
(Bendat and Piersol, 1986; Dalhaus, 2000)



Partial Directed Coherence

$$\pi_{ij}^2(f) = \frac{|A_{ij}(f)|^2}{\sum_{k=1}^{M} |A_{kj}(f)|^2}$$

(Baccalá and Sameshima, 2001)

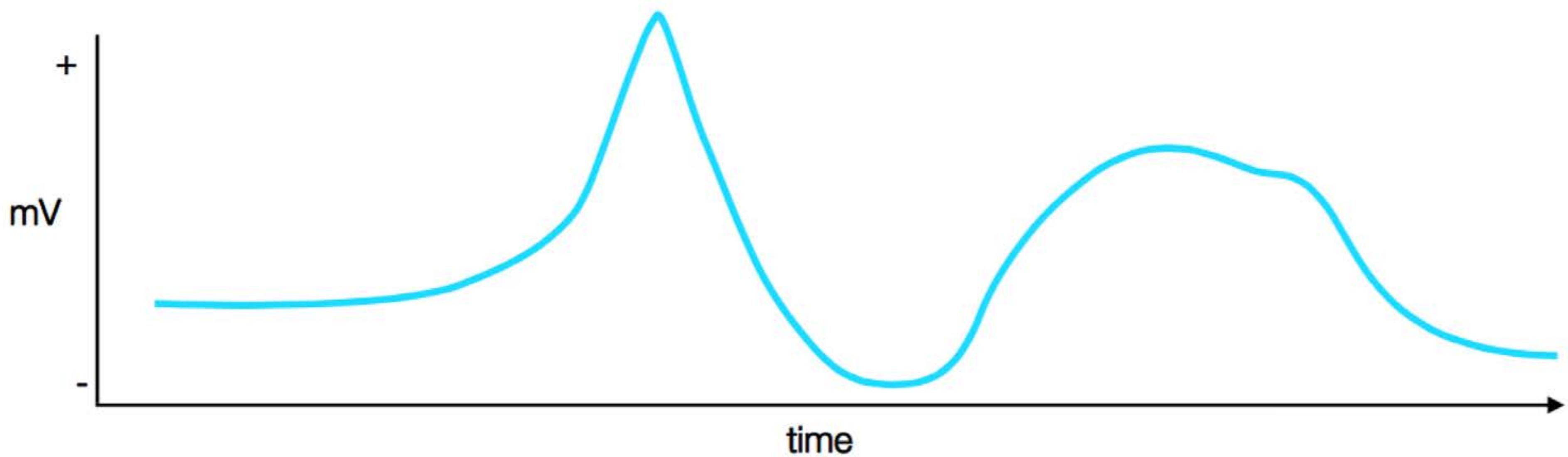


Time-Frequency GC

- Brain network dynamics often change rapidly with time
  - event-related responses
  - transient network changes during sequential information processing
- Electrophysiological processes often exhibit oscillatory phenomena, making them well-suited for frequency-domain analysis

Adapting to Non-Stationarity

- The brain is a **dynamic system** and measured brain activity and coupling can change rapidly with time (non-stationarity)
  - event-related perturbations (ERSP, ERP, etc)
  - structural changes due to learning/feedback
- How can we adapt to non-stationarity?

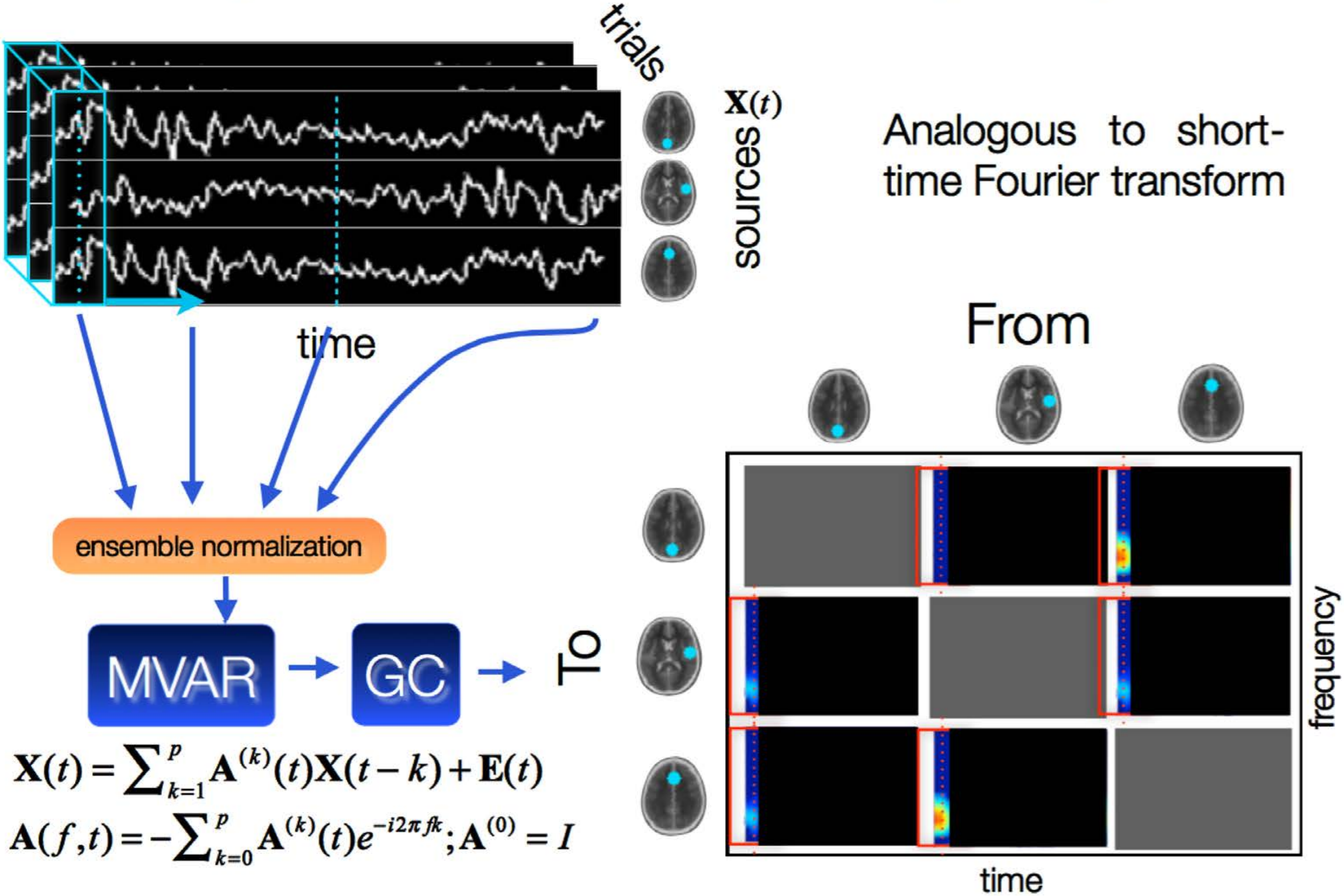


Adapting to Non-Stationarity

- **Many ways to do adaptive VAR estimation**
  - Segmentation-based adaptive VAR estimation
  - Factorization of time-varying spectral density matrices (e.g. from STFTs, Wavelets, etc)
  - State-Space Modeling
  - ...

Segmentation-based VAR

(Jansen et al., 1981; Florian and Pfurtscheller, 1995; Ding et al, 2000)

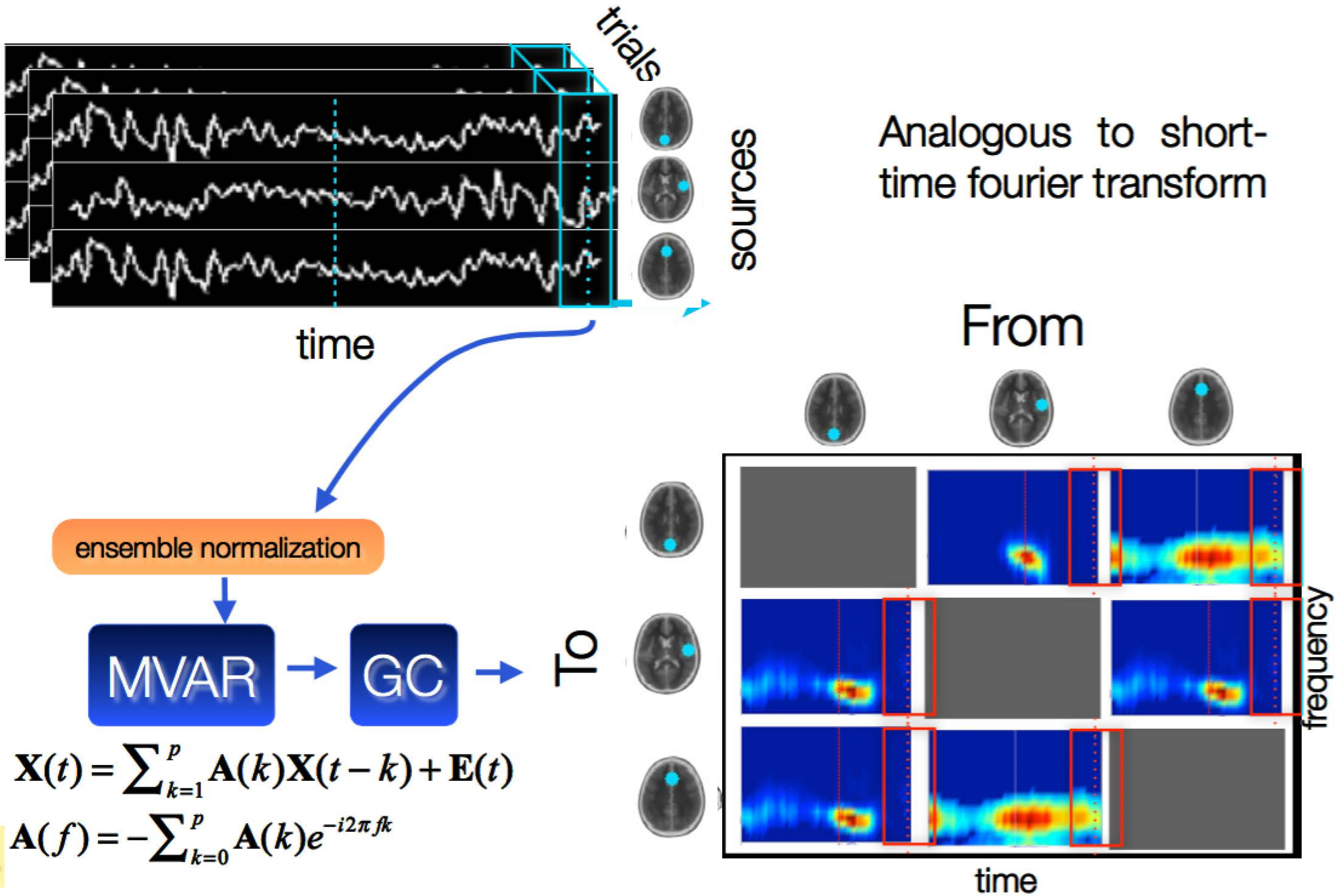


Analogous to short-time Fourier transform

$$\mathbf{X}(t) = \sum_{k=1}^{p} \mathbf{A}^{(k)}(t)\mathbf{X}(t-k) + \mathbf{E}(t)$$

$$\mathbf{A}(f,t) = -\sum_{k=0}^{p} \mathbf{A}^{(k)}(t)e^{-i2\pi fk}; \mathbf{A}^{(0)} = I$$

Analogous to short-time fourier transform



$$\mathbf{X}(t) = \sum_{k=1}^{p} \mathbf{A}(k)\mathbf{X}(t-k) + \mathbf{E}(t)$$

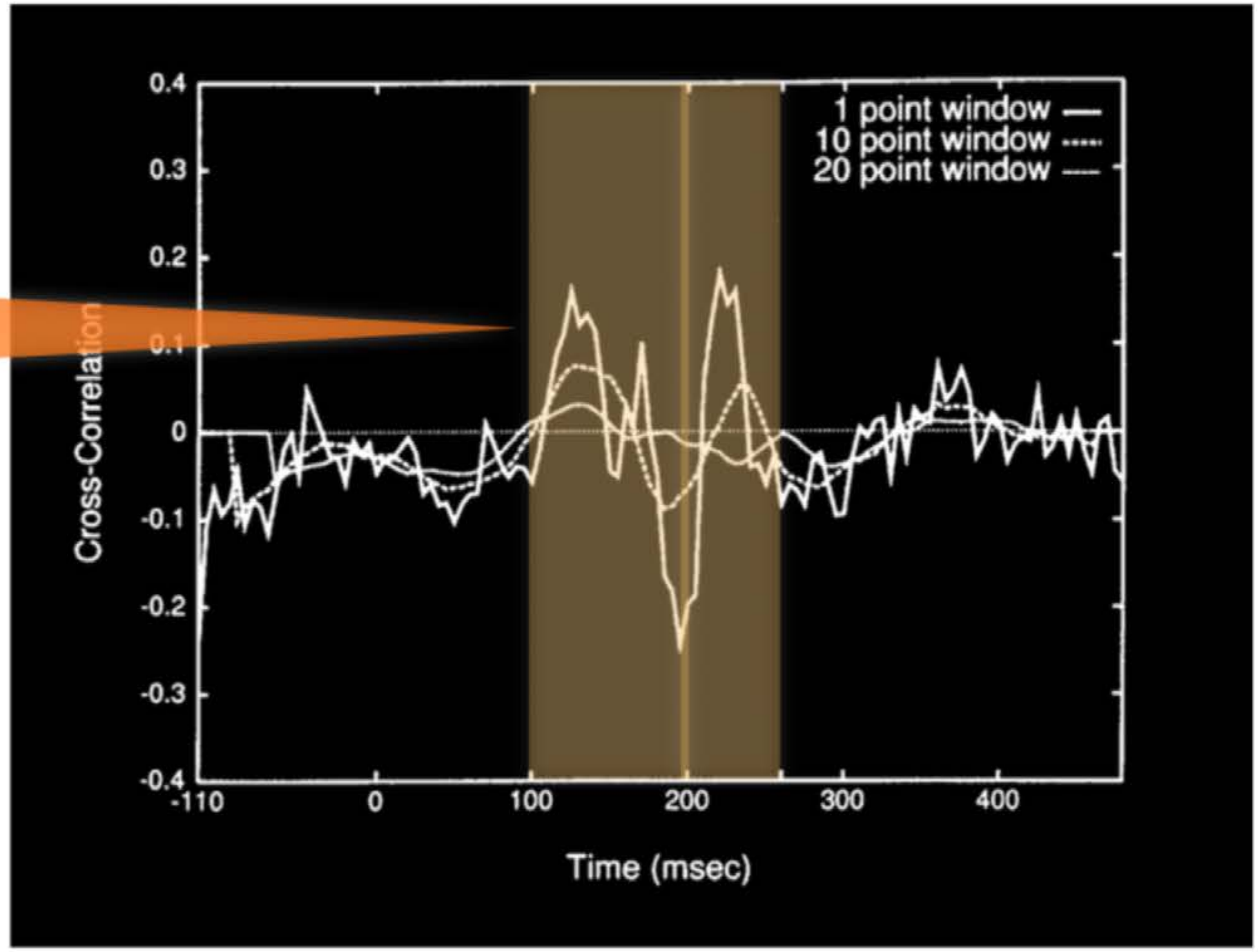
$$\mathbf{A}(f) = -\sum_{k=0}^{p} \mathbf{A}(k) e^{-i2\pi fk}$$

Important Choices

- Model Order
  - Determines complexity of spectrum you can model
  - Larger orders need more data
- Window Length
  - Window must be long enough to contain sufficient data for your chosen model order
  - Must be long enough to encompass the time-scale of interactions
  - Yet not too long as to smear temporal dynamics or include non-stationary data

Consideration: Local Stationarity

Too-large windows may not be locally-stationary



Consideration: Sufficient data

M = number of variables

p = model order

N_{tr} = number of trials

W = length of each window (sample points)

We have M^2p model coefficients to estimate. This requires a minimum of M^2p independent samples.

So we have the constraint $M^2p <= N_{tr} W$.

In practice, however, a better heuristic is $M^2p <= (1/10)N_{tr} W$.

Or: $\mathbf{W >= 10(M^2p/N_{tr})}$

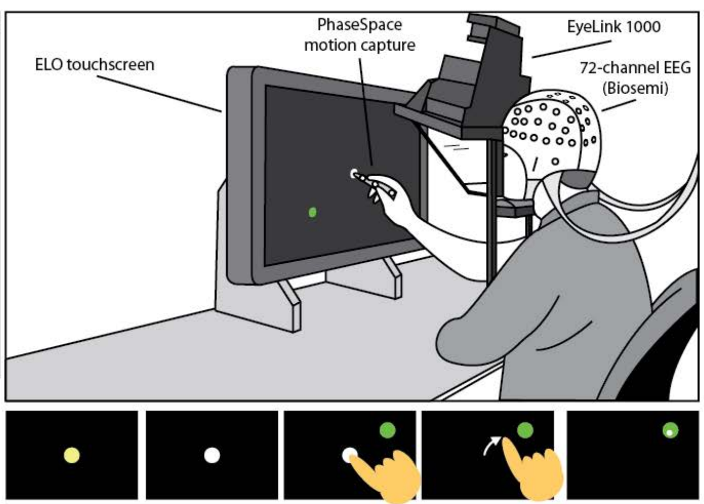
10x more data points than parameters to estimate

SIFT will let you know if your window length is not optimal

Network causal information flow during motor planning and execution

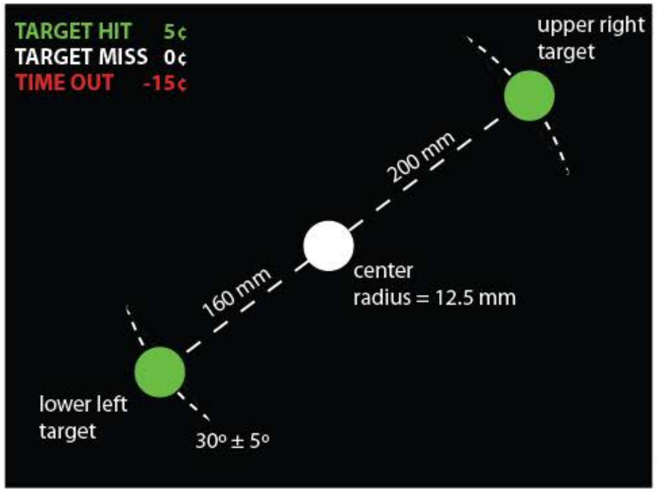
John R. Iversen, Alejandro Ojeda, Tim Mullen, Markus Plank, Joseph Snider, Gert Cauwenberghs, Howard Poizner

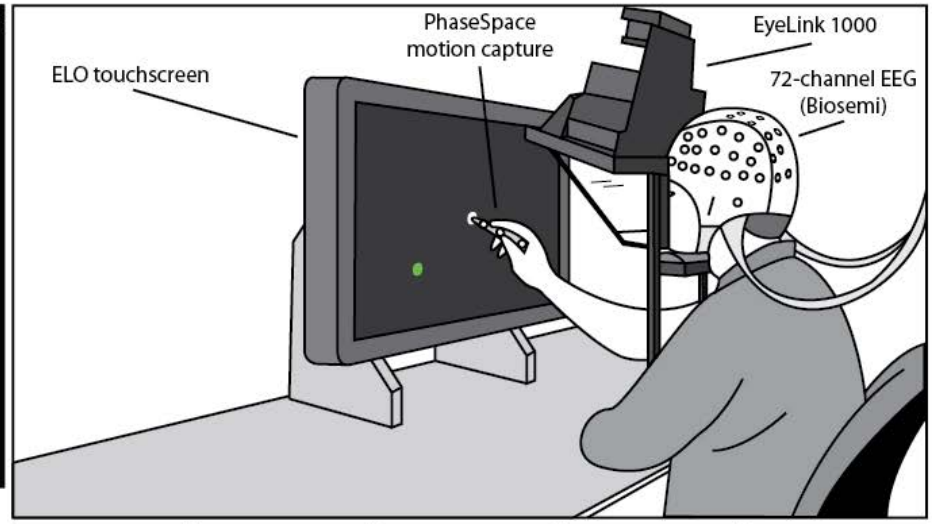
Institute for Neural Computation
Swartz Center for Computational Neuroscience
University of California, San Diego
EMBC 2014

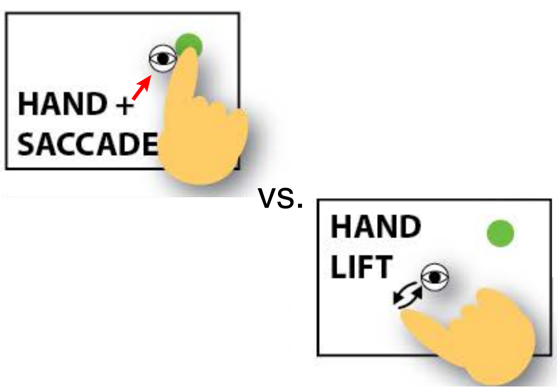


How does brain plan visually guided movements?

- Pointing Task (Park, et al. 2014, *IEEE Trans Neural Syst Rehabil Eng)*







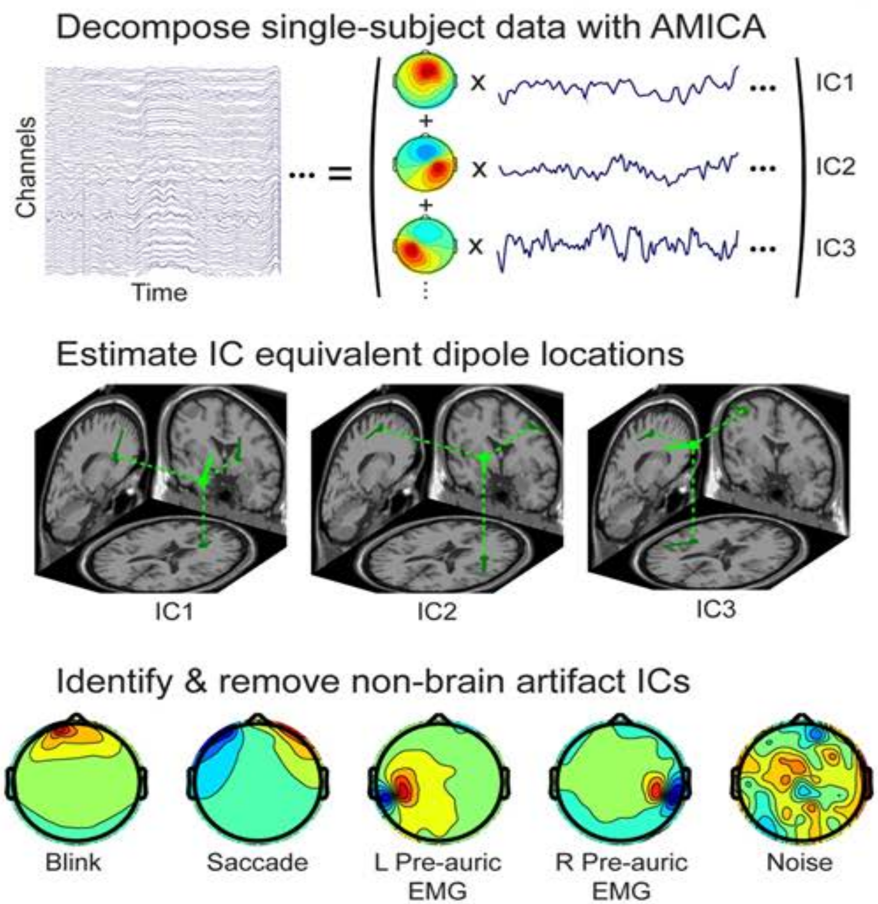
1	2	3	4	5
hold	target onset hold	center offset	eye/hand initiation	eye/hand arrival, feedback
500 - 700	500 - 700	< time window (e.g., 450 ms)		

N=10 (right-handed, mean age=21)
70 channel EEG (Biosemi)
512 Hz; 128Hz for connectivity

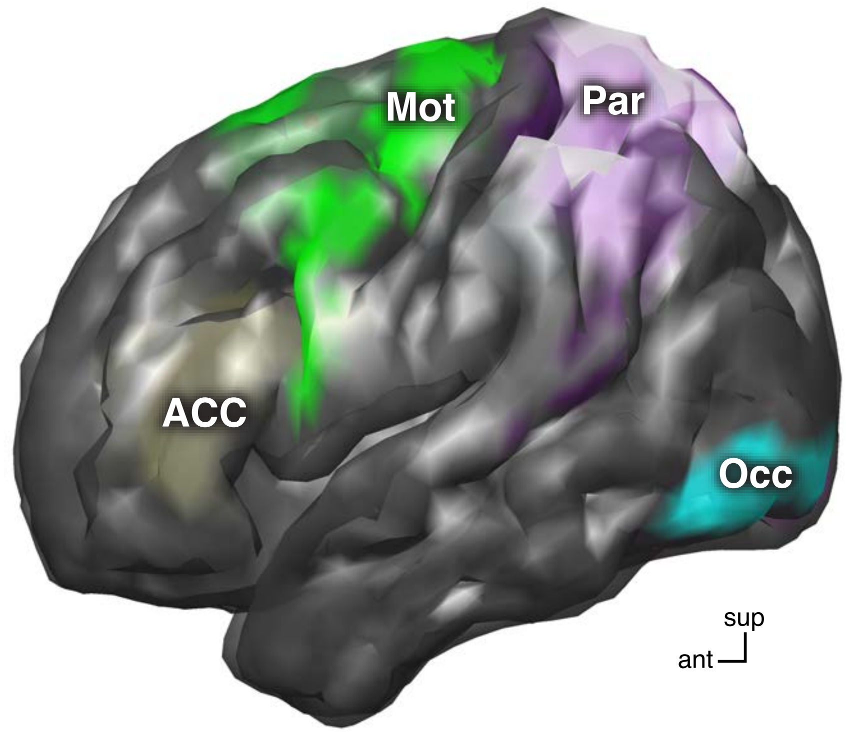
Planning | Execution ← Analysis Window
0 0.5 1.0
Time (s)

ICA source space analysis

Independent Component Analysis



Cortical ROIs



Group SIFT: Project ICs onto cortical surface using LORETA; extract ROI time series. Advantage: Same ROIs for all subjects enables statistical comparison. (*Use BCILAB srcpot*)

Core Analysis Methods I

- Segmentation–based MVAR

$$\mathbf{X}(t) = \sum_{k=1}^{p} \mathbf{A}^{(k)}(t)\mathbf{X}(t-k) + \mathbf{E}(t)$$

$$\mathbf{A}(f) = -\sum_{k=0}^{p} \mathbf{A}^{(k)} e^{-i2\pi fk}; \mathbf{A}^{(0)} = I$$

$$\mathbf{A}(f)^{-1} = \mathbf{H}(f)$$

$$A(t) = \arg\min_{\hat{A}} \left(\overbrace{\left\| Y - Z\tilde{A} \right\|_2^2}^{\text{prediction error}} + \underbrace{\boxed{\lambda}}_{\text{regularization}} \underbrace{\sum_{ij} \overbrace{\left\| \tilde{A}_{ij}^{(1)}, \ldots, \tilde{A}_{ij}^{(p)} \right\|_2}^{\text{smoothness (L2) (preserves spectrum)}}}_{\text{group sparsity (L1)}} \right)$$

Core Analysis Methods II

- Time-varying SdDTF ("short-time direct directed transfer function")

- Directed measure of direct (unmediated) causal flow between ROIs

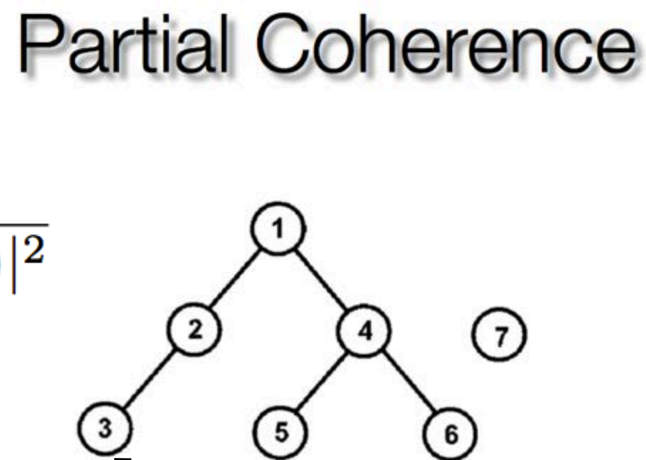
- Combines DTF and partial coherence; windowed (0.5s, 30ms).

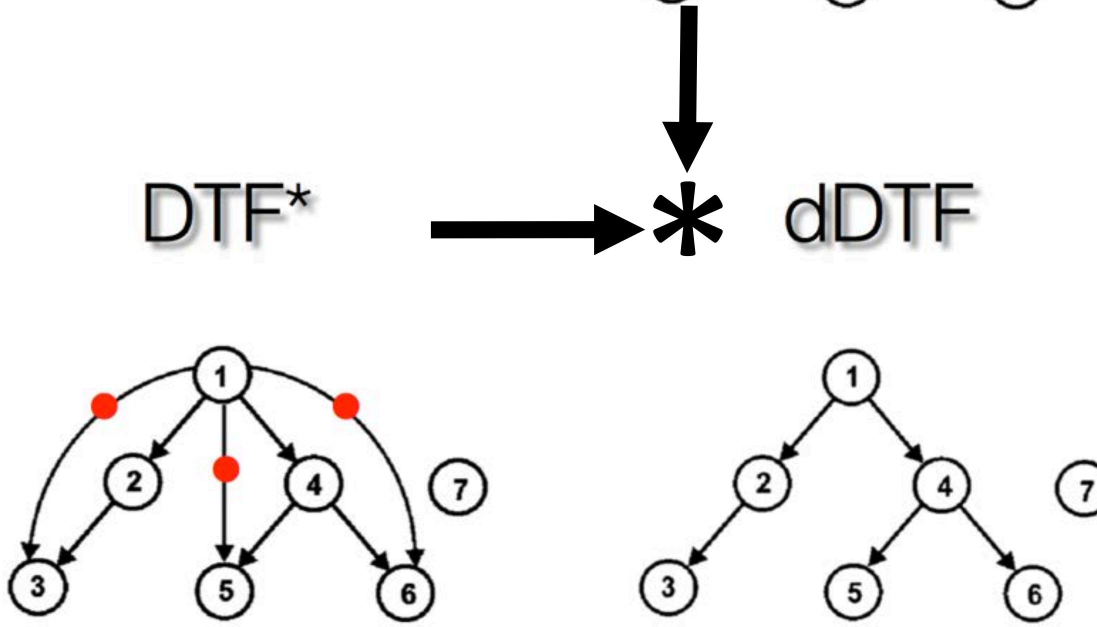
$$\eta_{ij}^2(f,t) = \frac{|H_{ij}(f,t)|^2 |P_{ij}(f,t)|^2}{\sum_{klf\tau} |H_{kl}(f,\tau)|^2 |P_{kl}(f,\tau)|^2}$$

(Korzeniewska, et al. 2008)

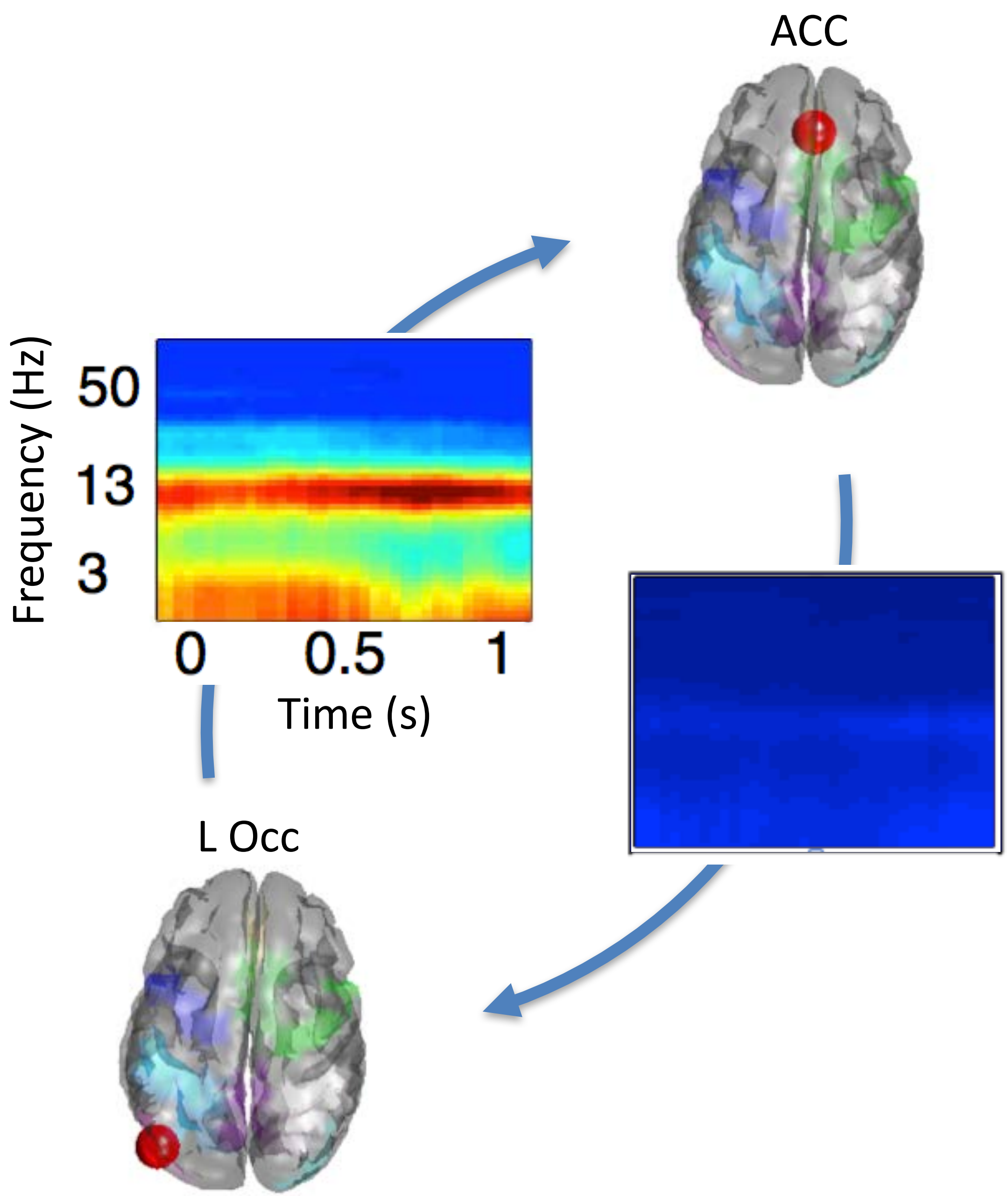
dDTF

$$\eta_{ij}^2(f,t) = \frac{|H_{ij}(f,t)|^2 |P_{ij}(f,t)|^2}{\sum_{klf\tau} |H_{kl}(f,\tau)|^2 |P_{kl}(f,\tau)|^2}$$





SIFT Analysis

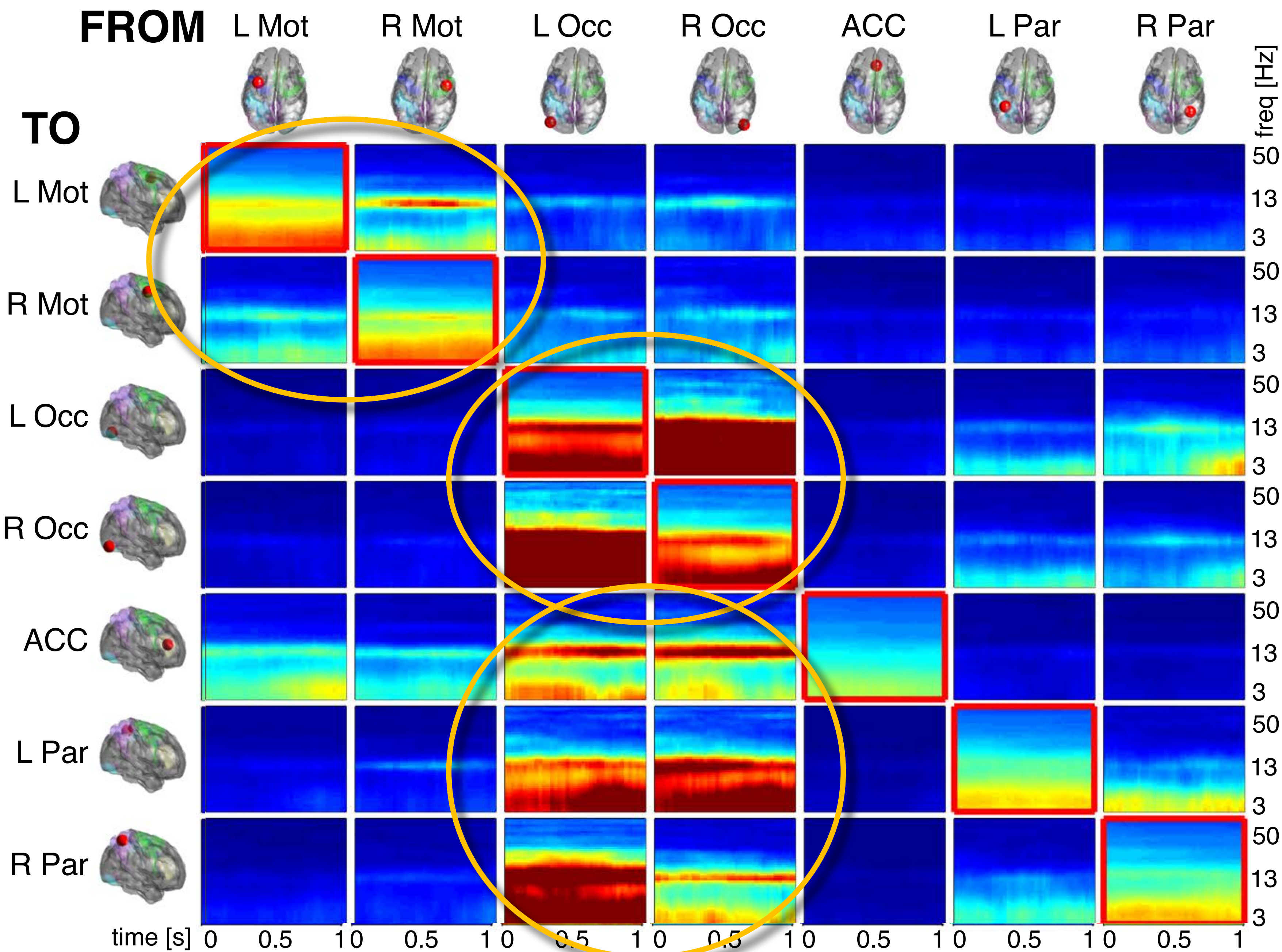


- Time-varying SdDTF

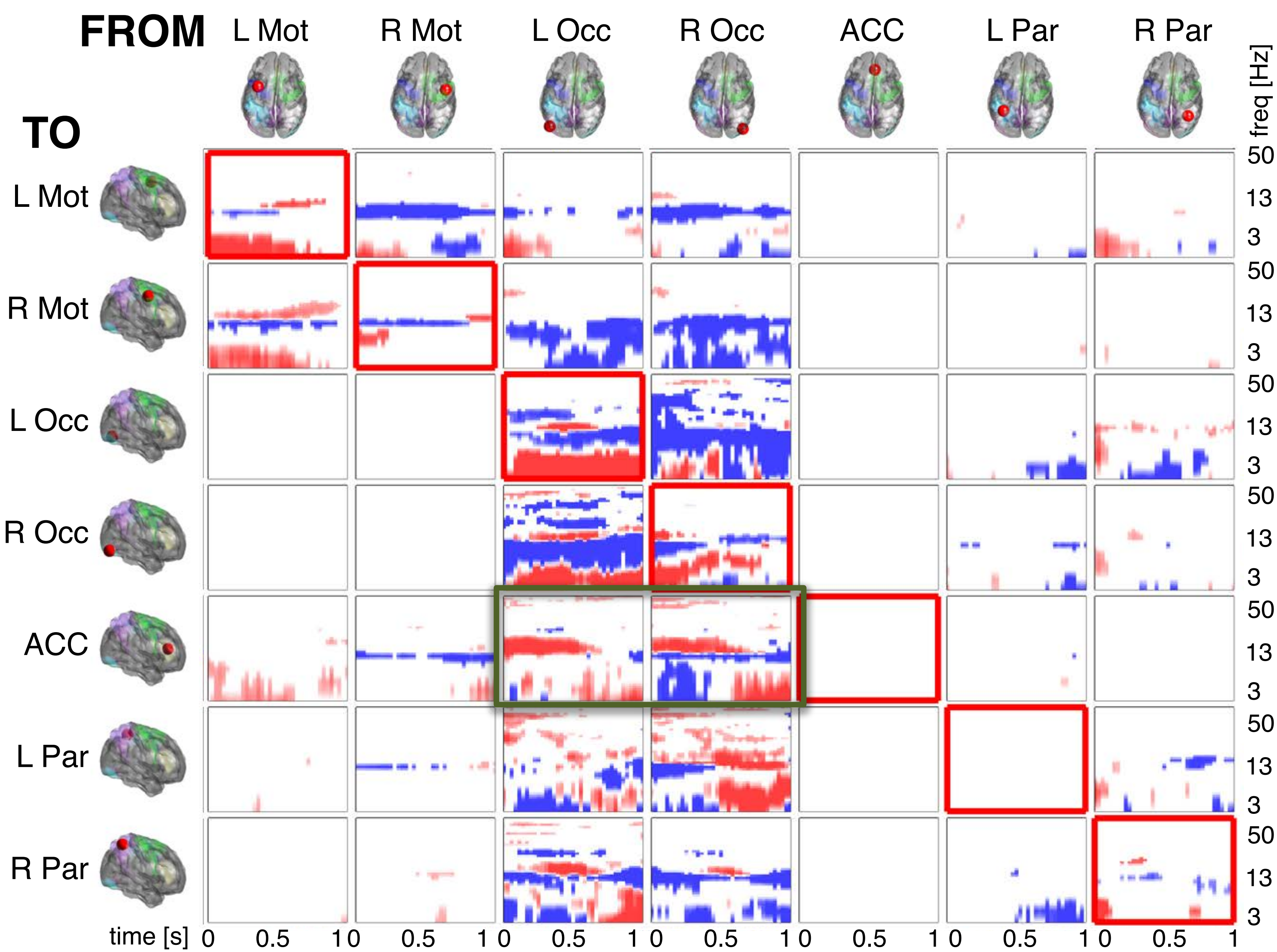
Directed measure of direct causal flow between ROIs

Averaged across subjects

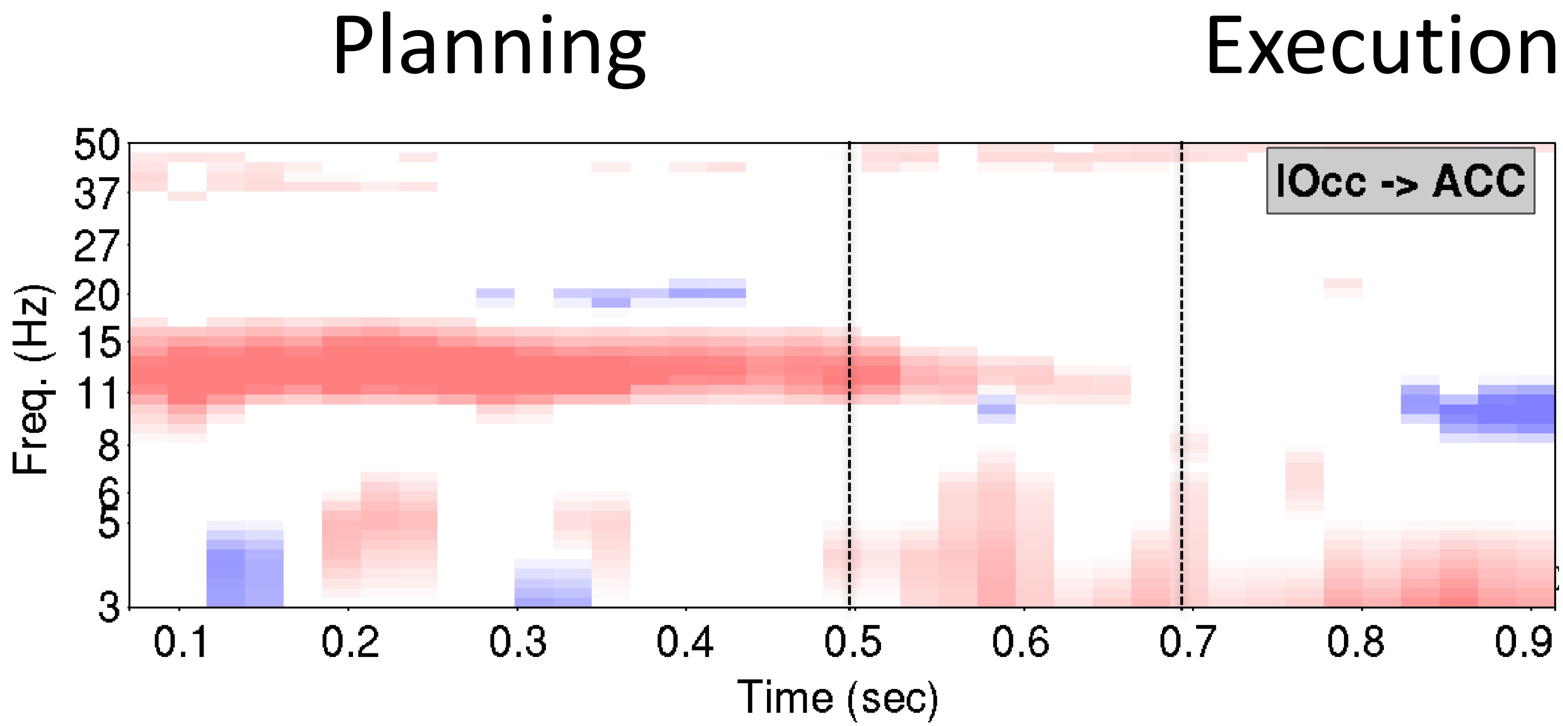
dDTF during reaching



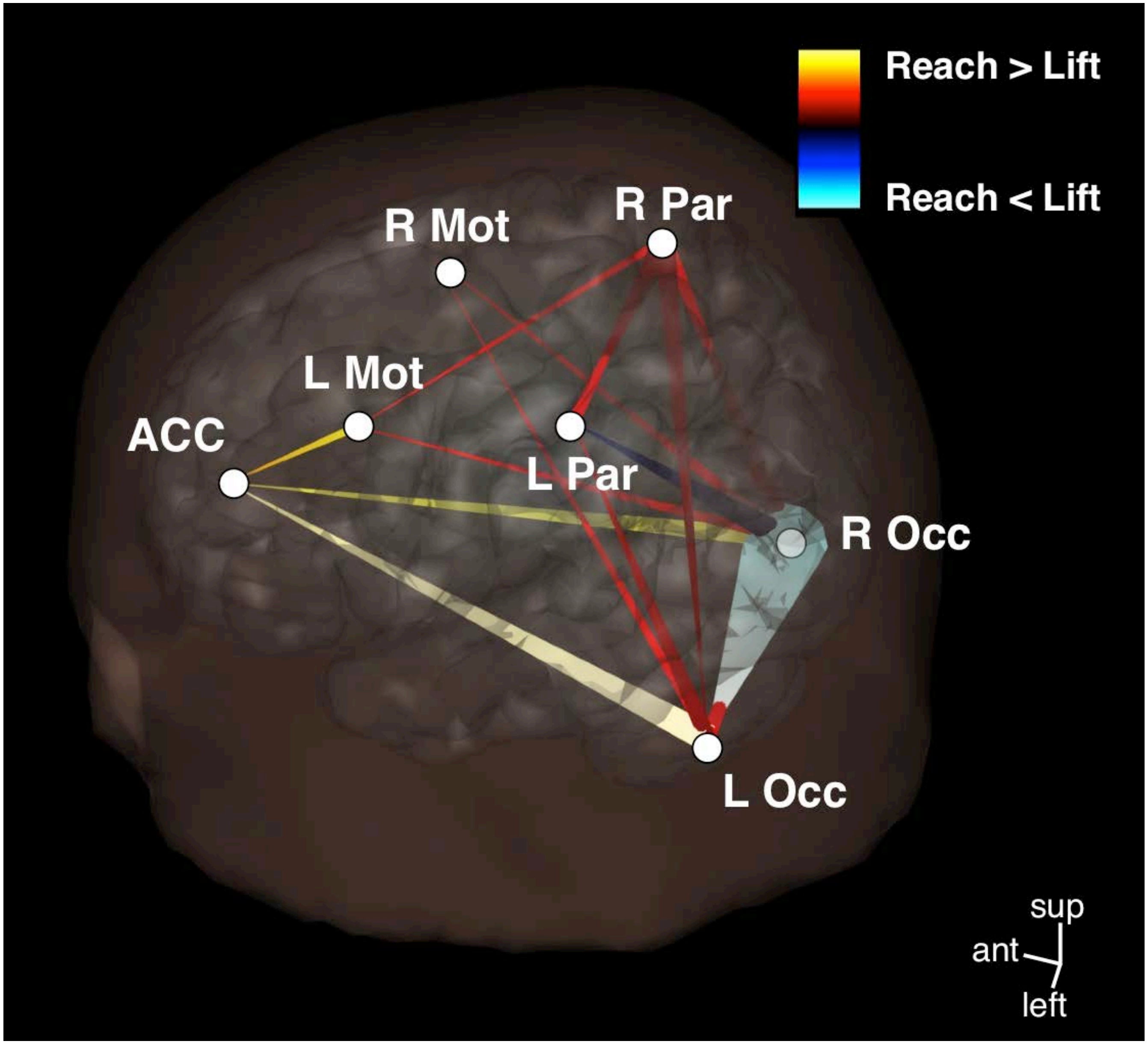
Changed causal flow during reaching



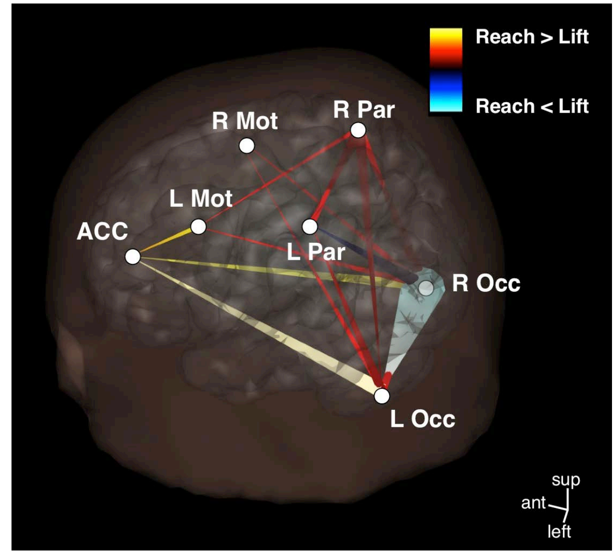
Occipital → ACC



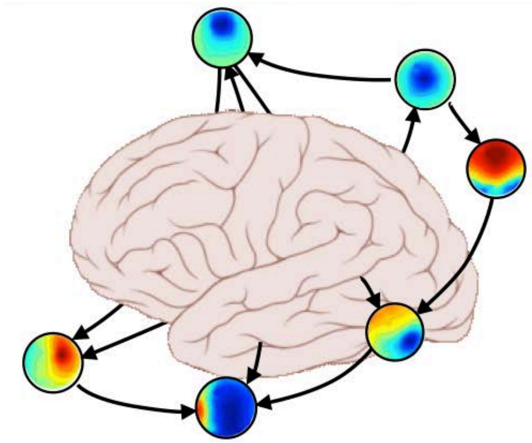
Greater causal flow during movement planning



Discussion



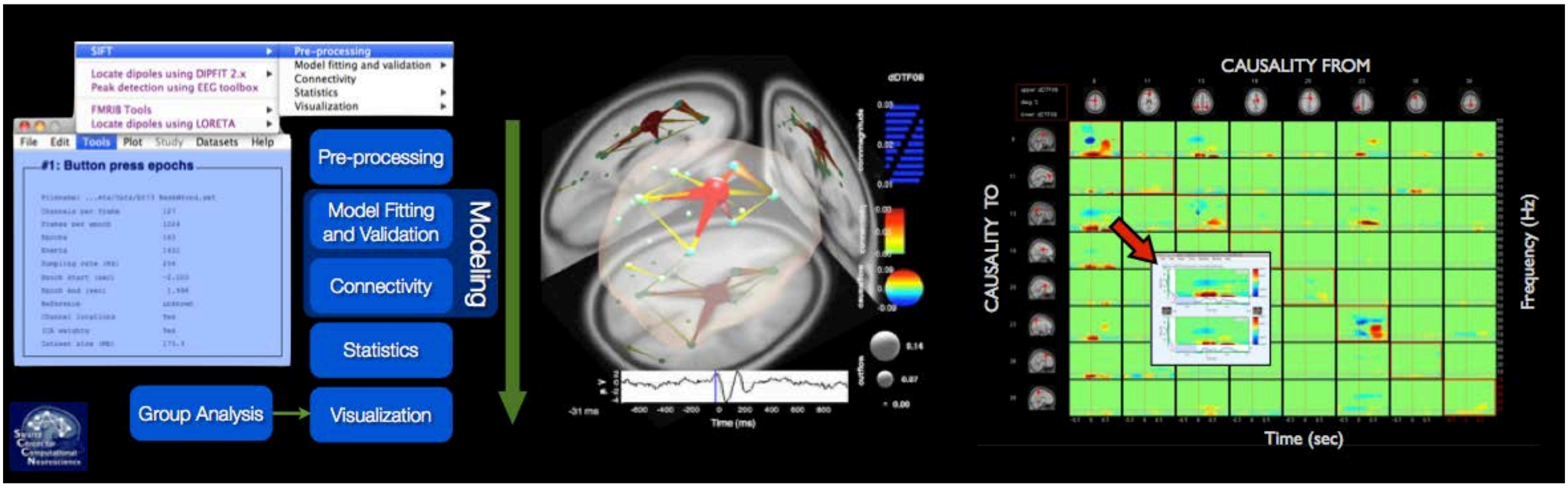
- SIFT is a capable toolkit for causal dynamical analysis at source level
- **Parietal** network expected for visually guided action (e.g. Heider, et al., 2010)
- **ACC** more strongly driven by Occipital & Motor. Locus for translation of intention into action (Paus, 2001; Srinivasan, et al. 2013). ACC drives SMA (not shown).
- Causal network results depend on the number of nodes
  - E.g. Occipital → ACC could be mediated by region not included in model
  - There will always be a tradeoff between network size and amount of data needed to fit the model.
  - Regularization



SIFT

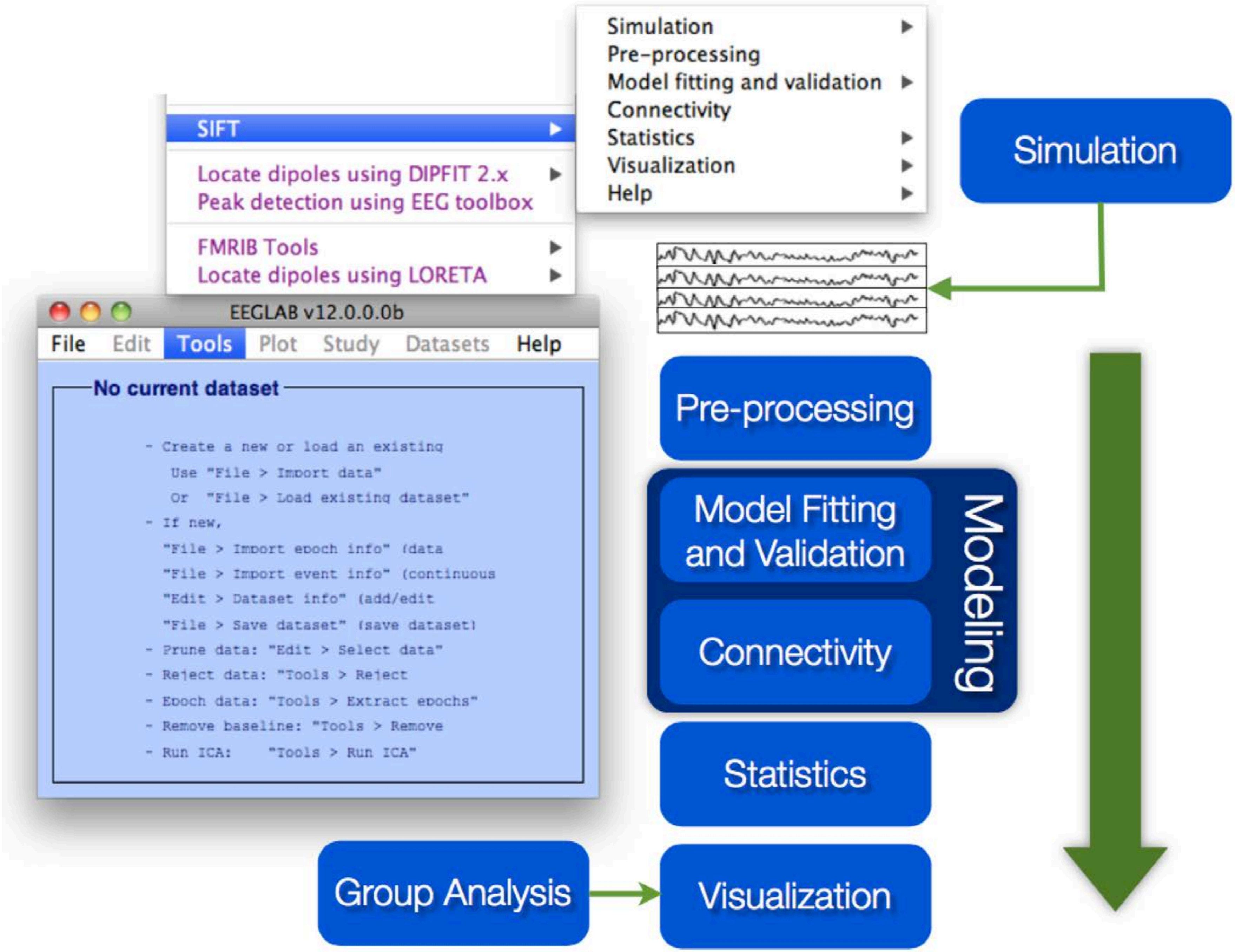
Source Information Flow Toolbox

http://sccn.ucsd.edu/wiki/SIFT
Mullen, et al, *Journal of Neuroscience Methods* (in prep, 2012)
Mullen, et al, *Society for Neuroscience*, 2010
Delorme, Mullen, Kothe et al, *Computational Intelligence and Neuroscience*, vol 12, 2011

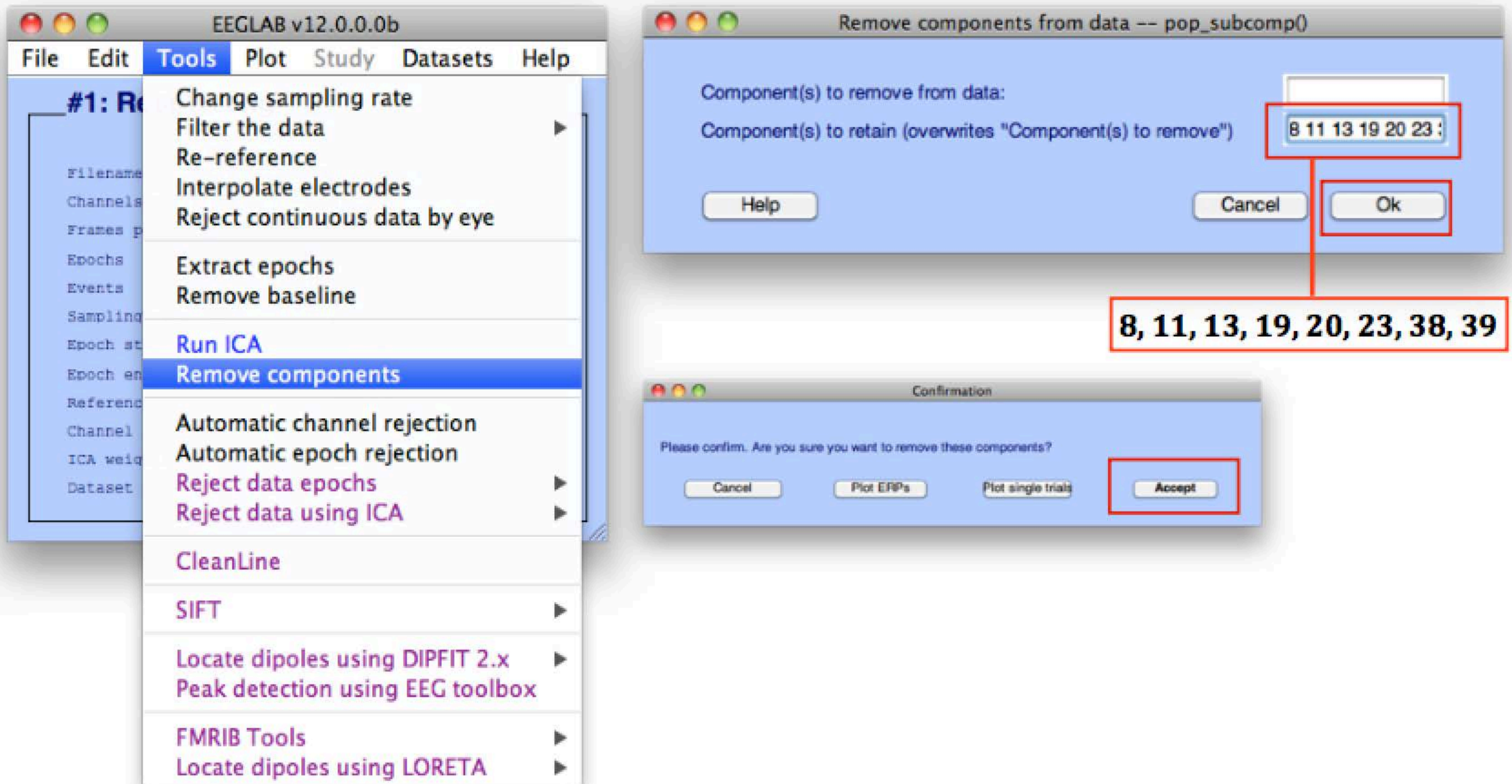


- A toolbox for (source-space) electrophysiological information flow and causality analysis (single- or multi-subject) integrated into the EEGLAB software environment.
- Emphasis on vector autoregression and time-frequency domain approaches
- Standard and novel interactive visualization methods for exploratory analysis of connectivity across time, frequency, and spatial location

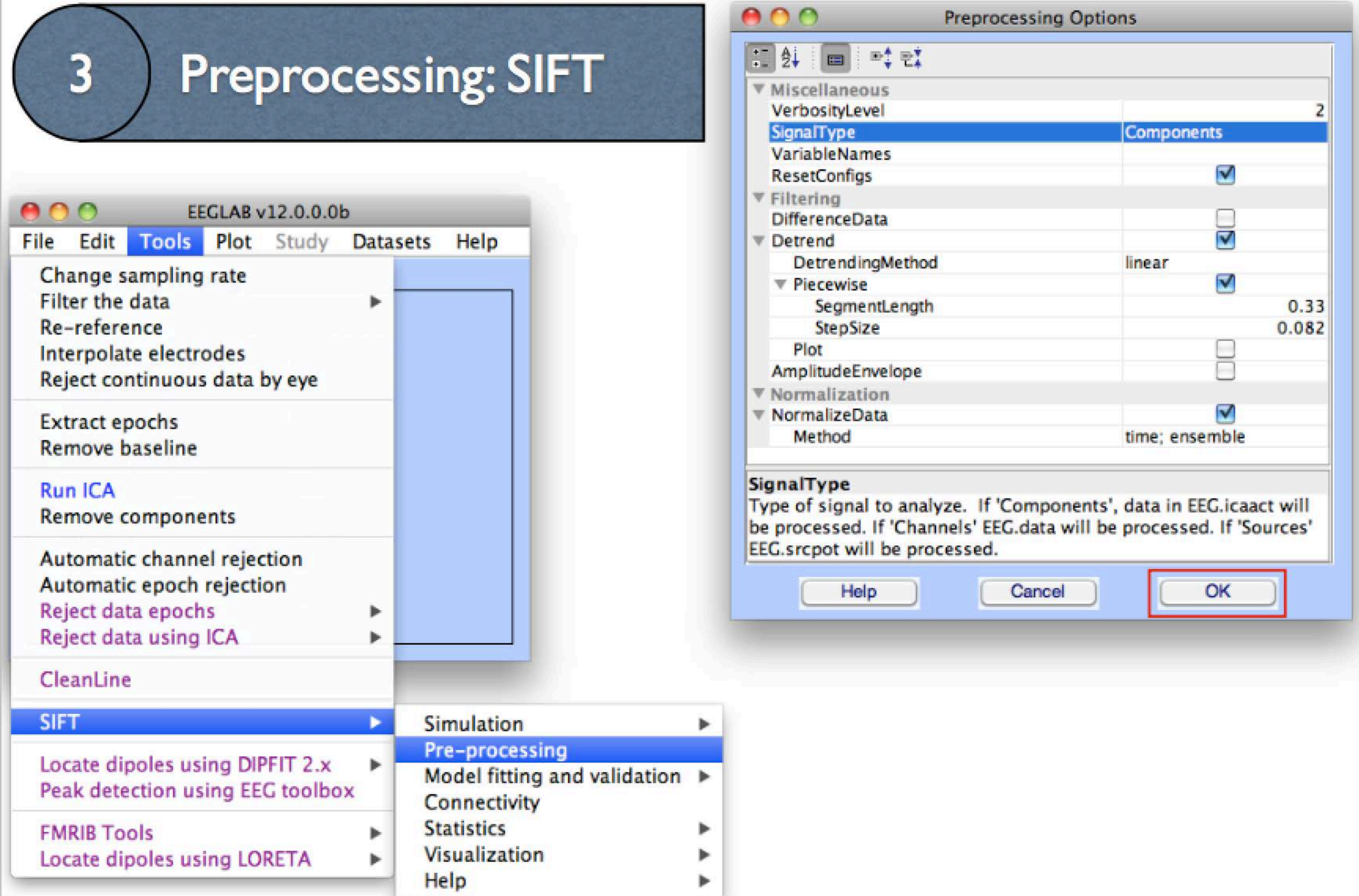
SIFT Workflow



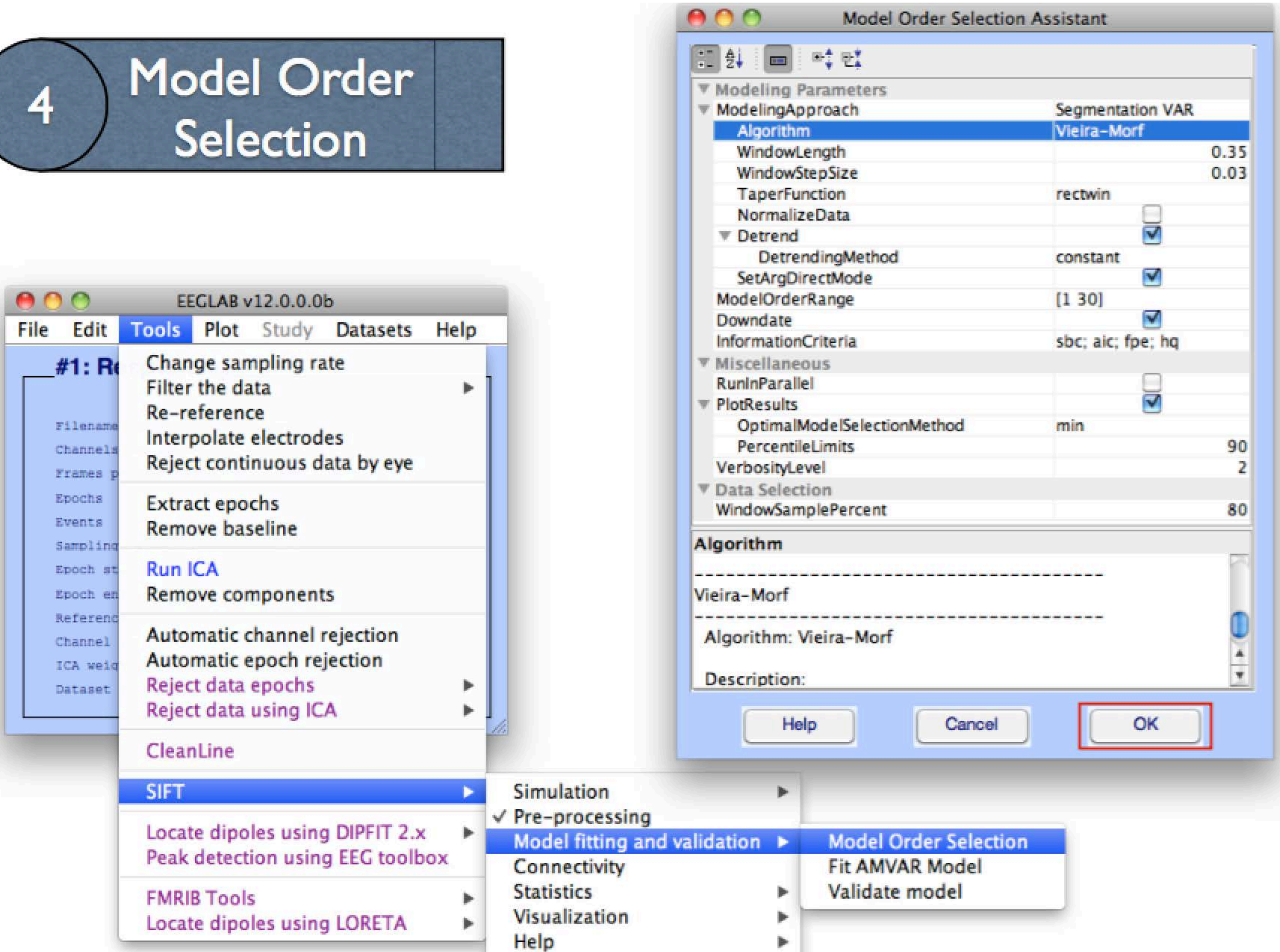
3 Preprocessing: Select Components



3 Preprocessing: SIFT



4 Model Order Selection



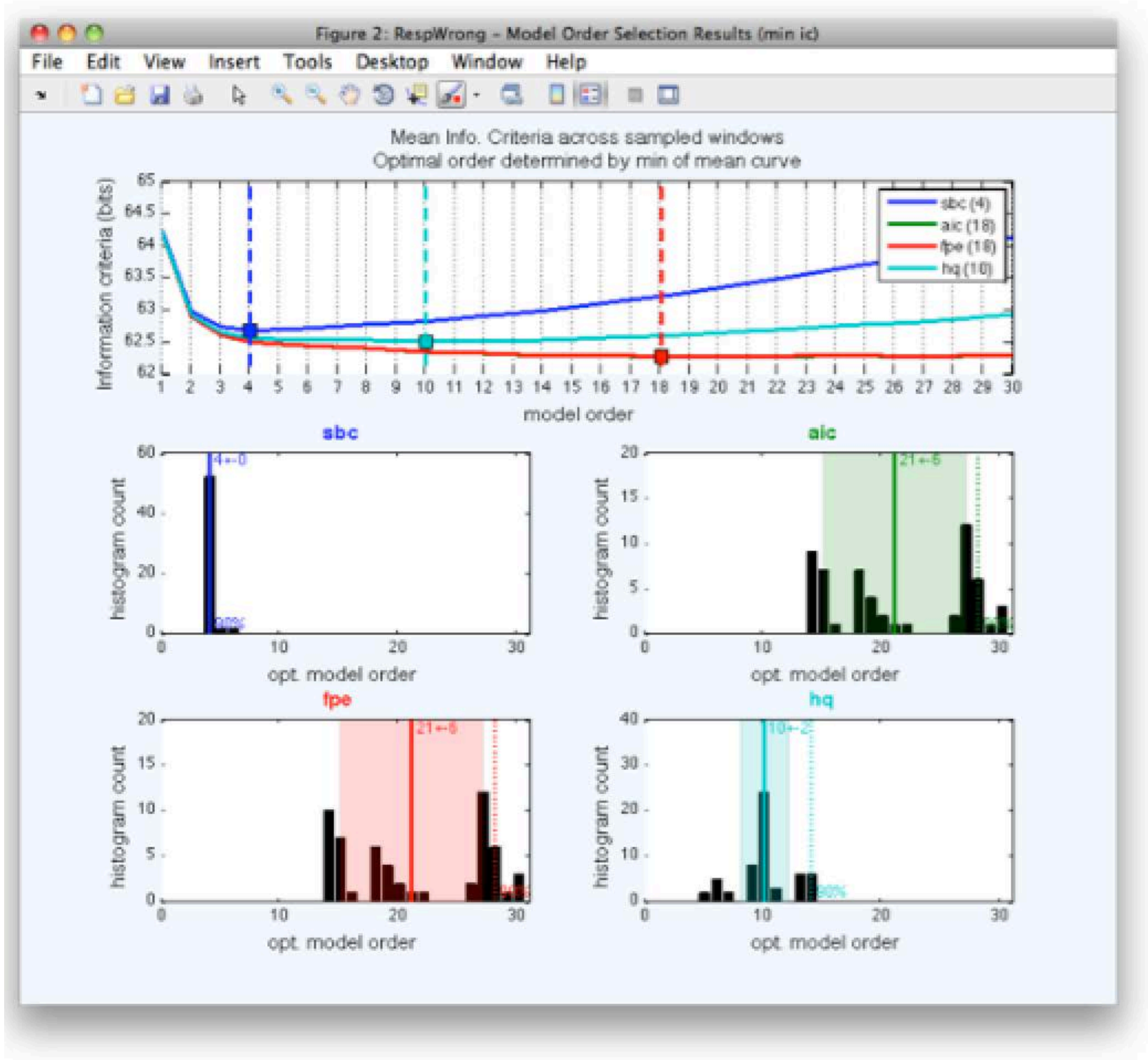
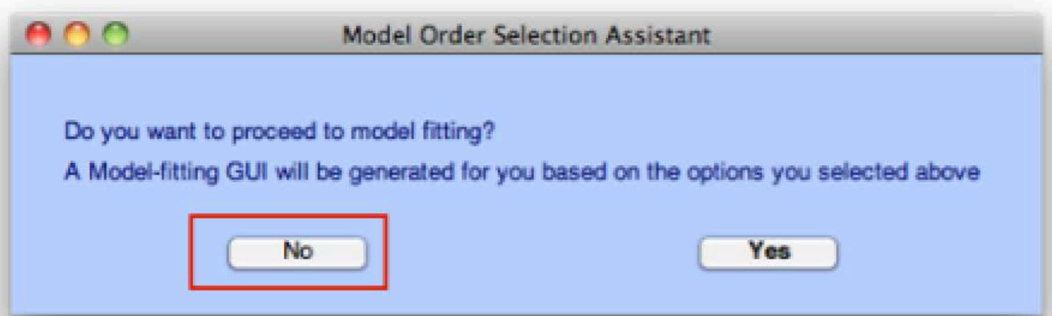
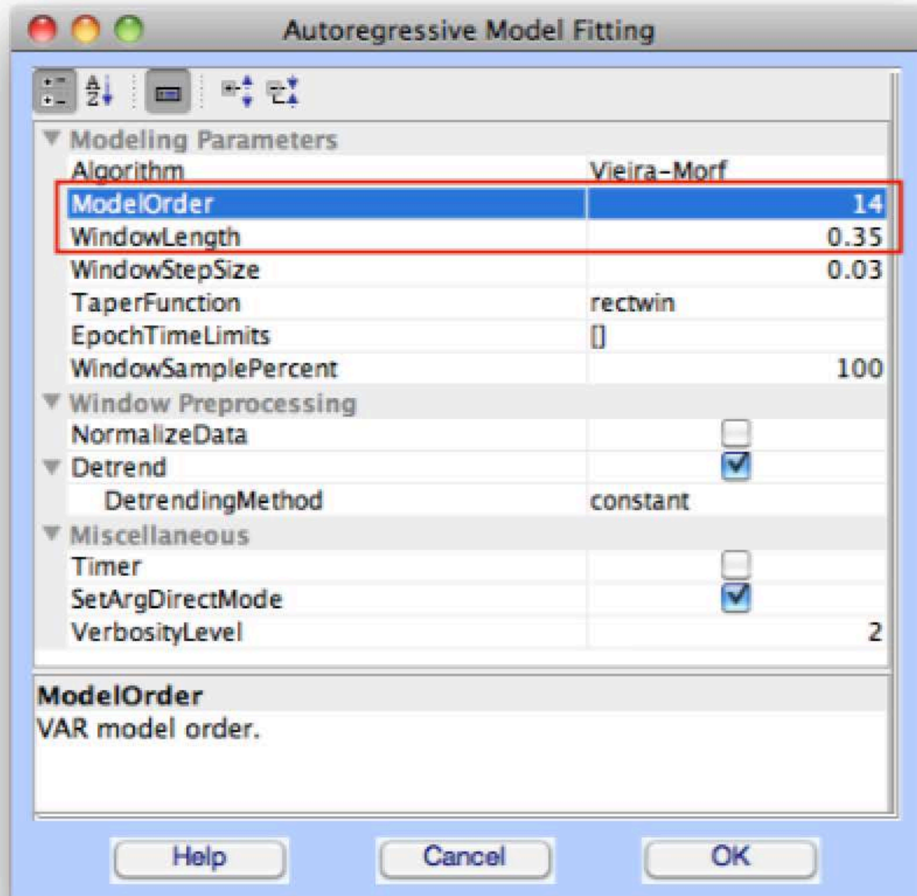
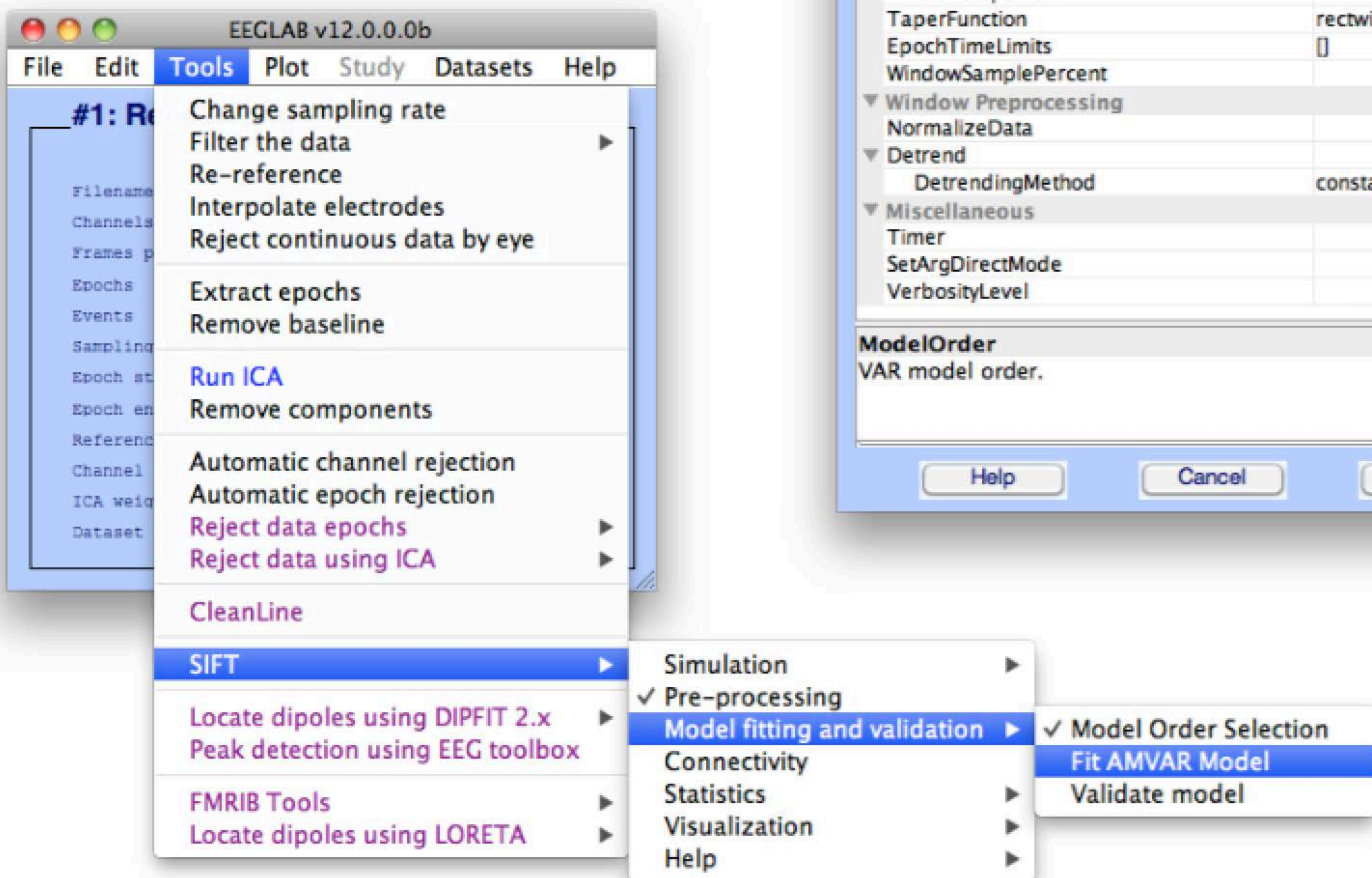


Figure 2: RespWrong - Model Order Selection Results (min ic)
File Edit View Insert Tools Desktop Window Help
Mean Info. Criteria across sampled windows
Optimal order determined by min of mean curve
Information criteria (bits)
model order
sbc (4)
aic (18)
fpe (18)
hq (10)
sbc
aic
fpe
hq
histogram count
opt. model order

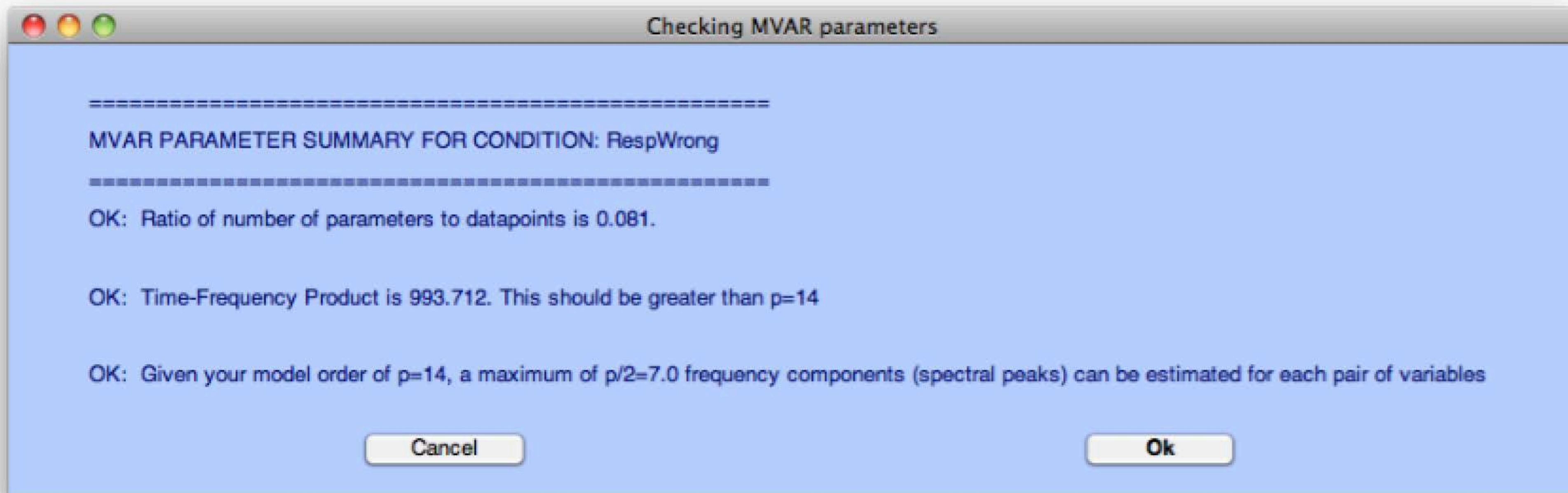

Model Order Selection Assistant
Do you want to proceed to model fitting?
A Model-fitting GUI will be generated for you based on the options you selected above
No
Yes

5 Model Fitting

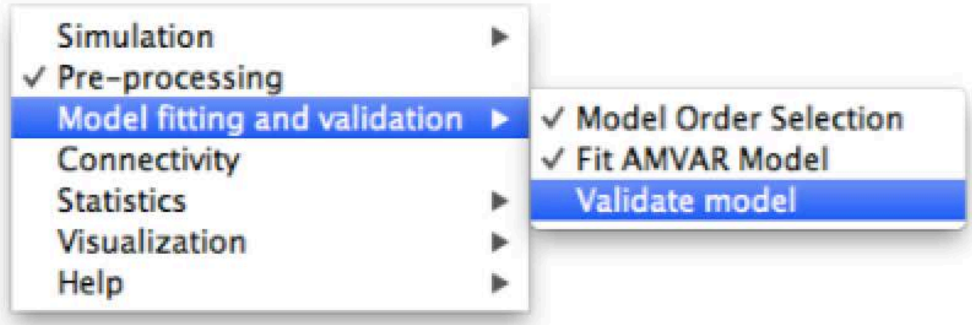


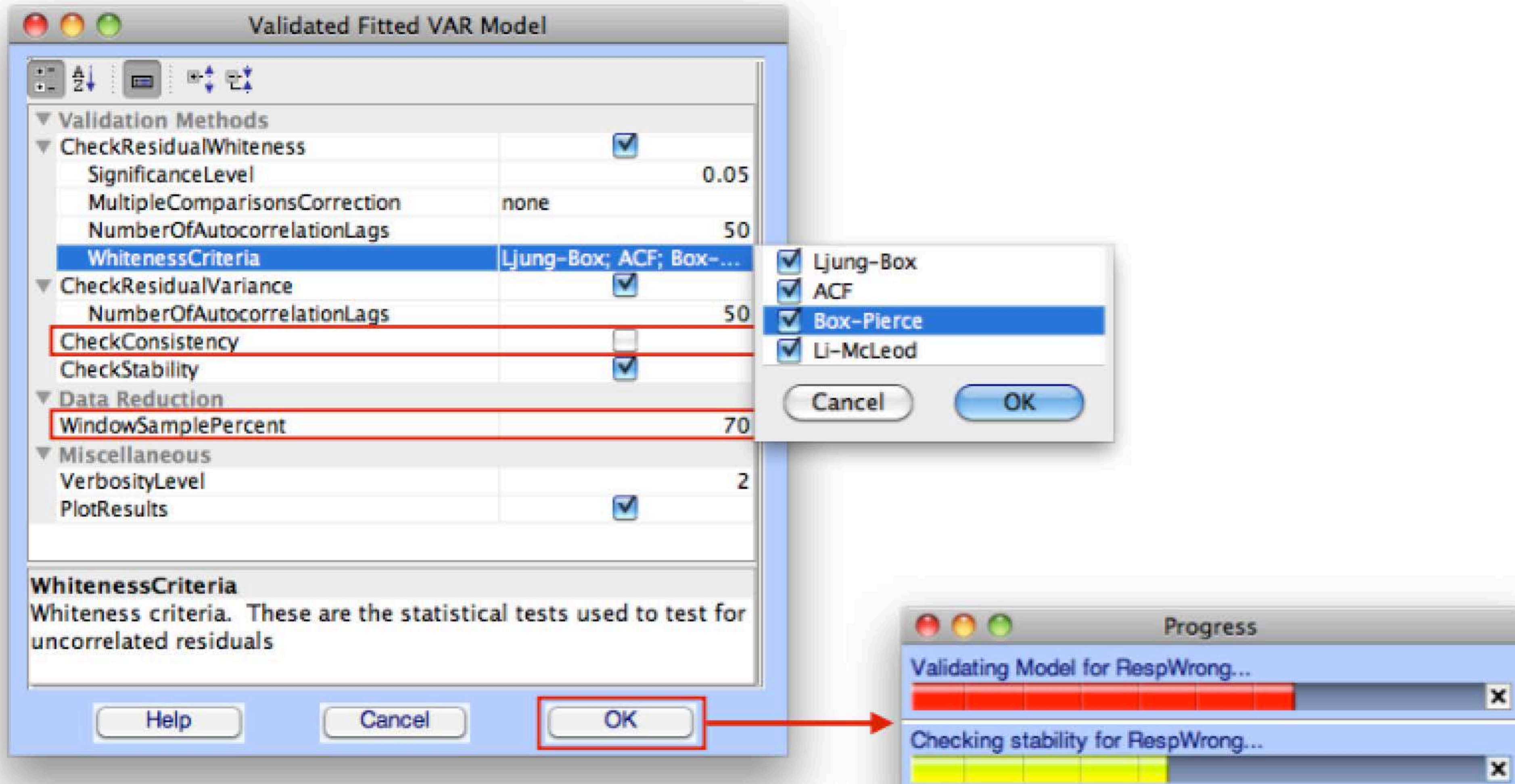


5 Model Fitting

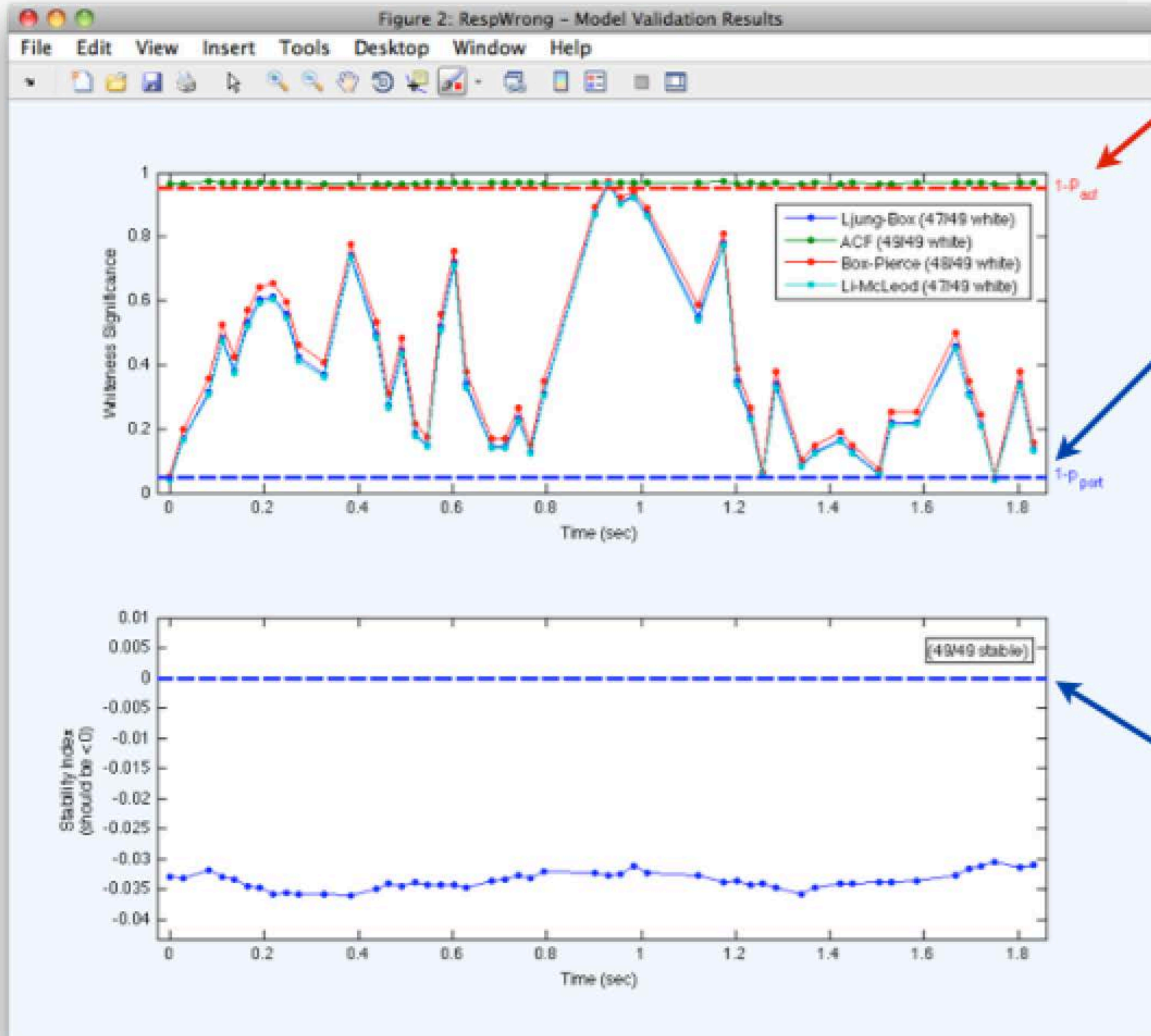


6 Model Validation





6 Model Validation

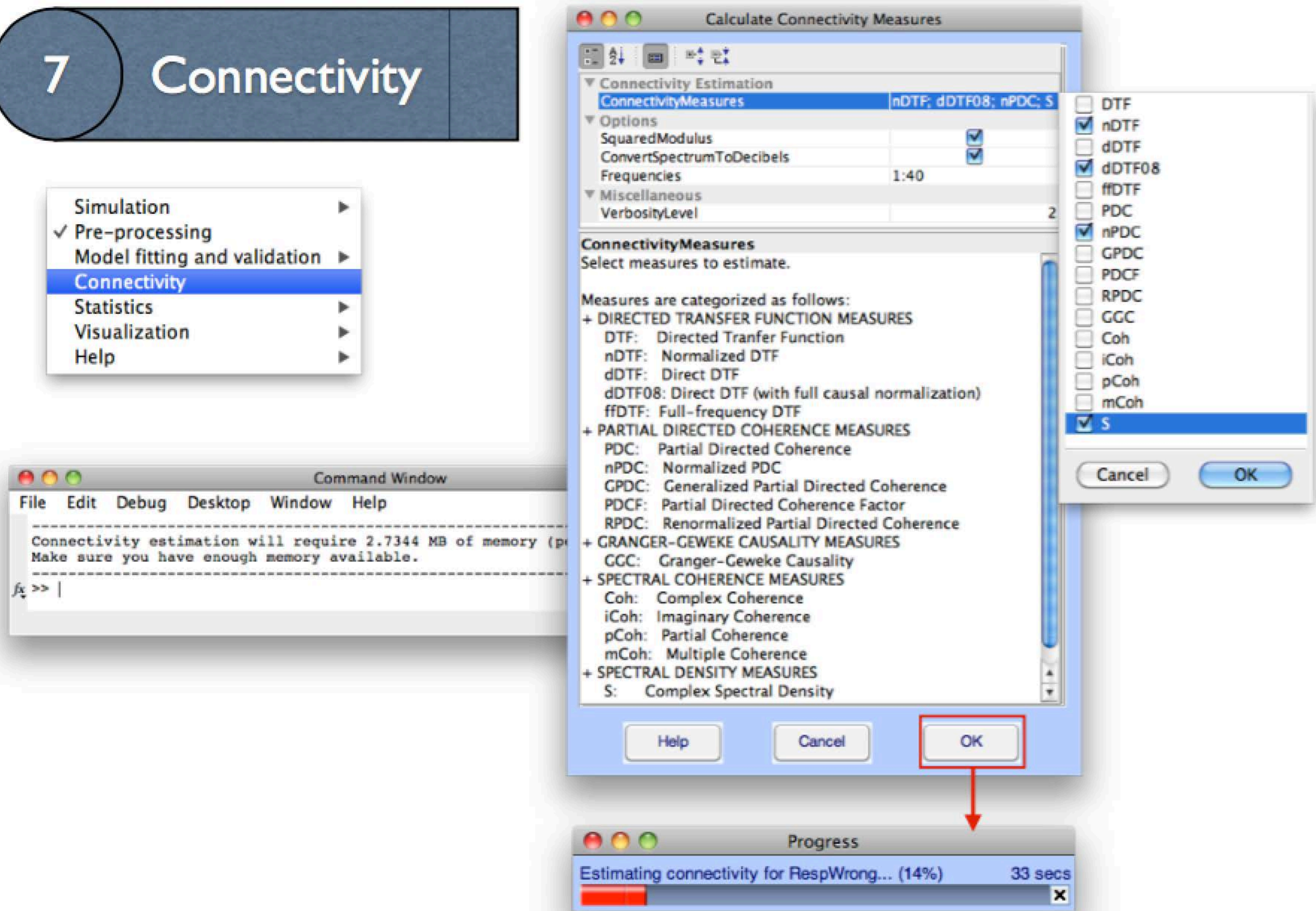


ACF statistic should be above this line

Portmanteau statistics should be above this line

Stability index should be < 0

7 Connectivity



8 Visualization: Time-Frequency Grid

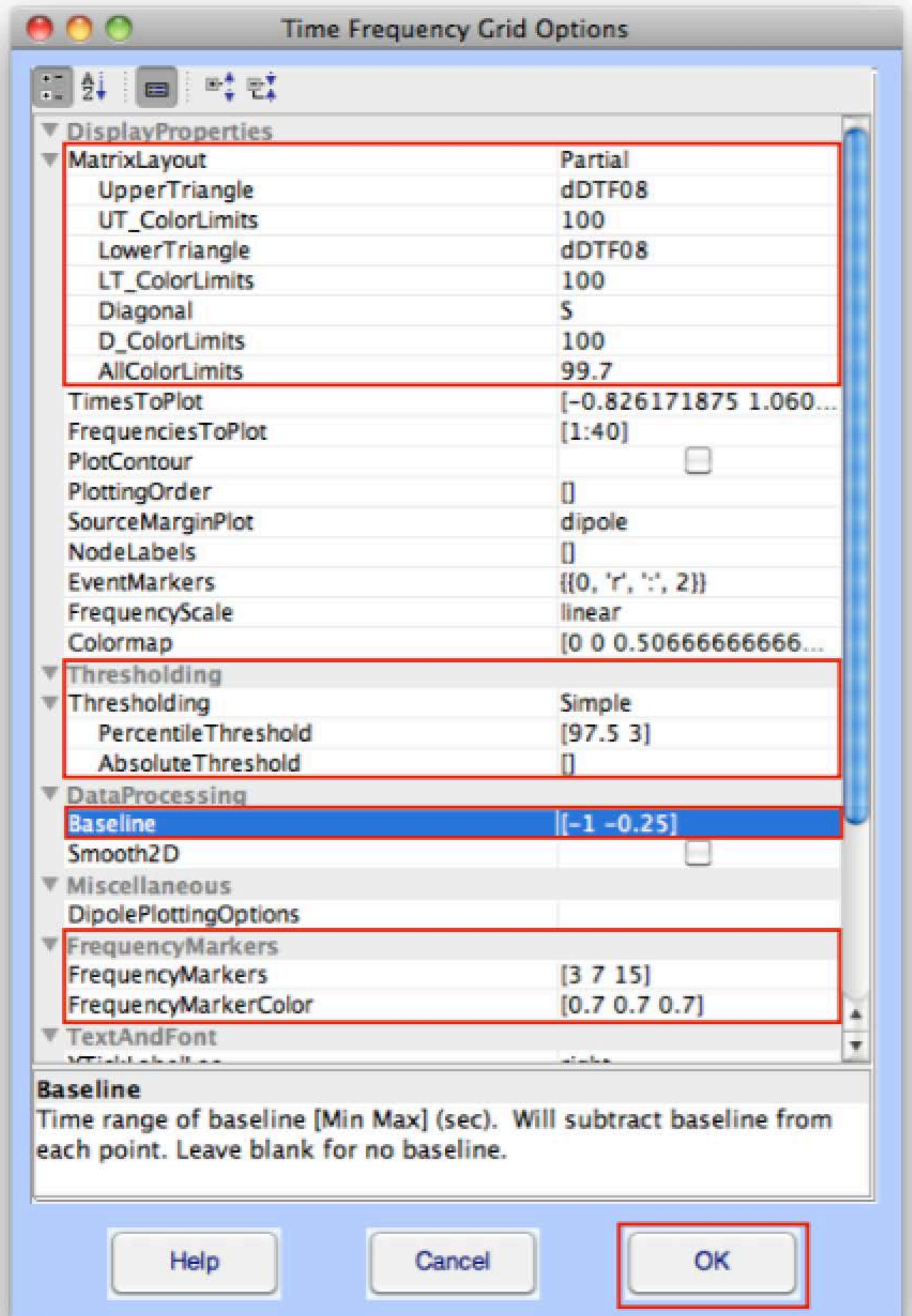
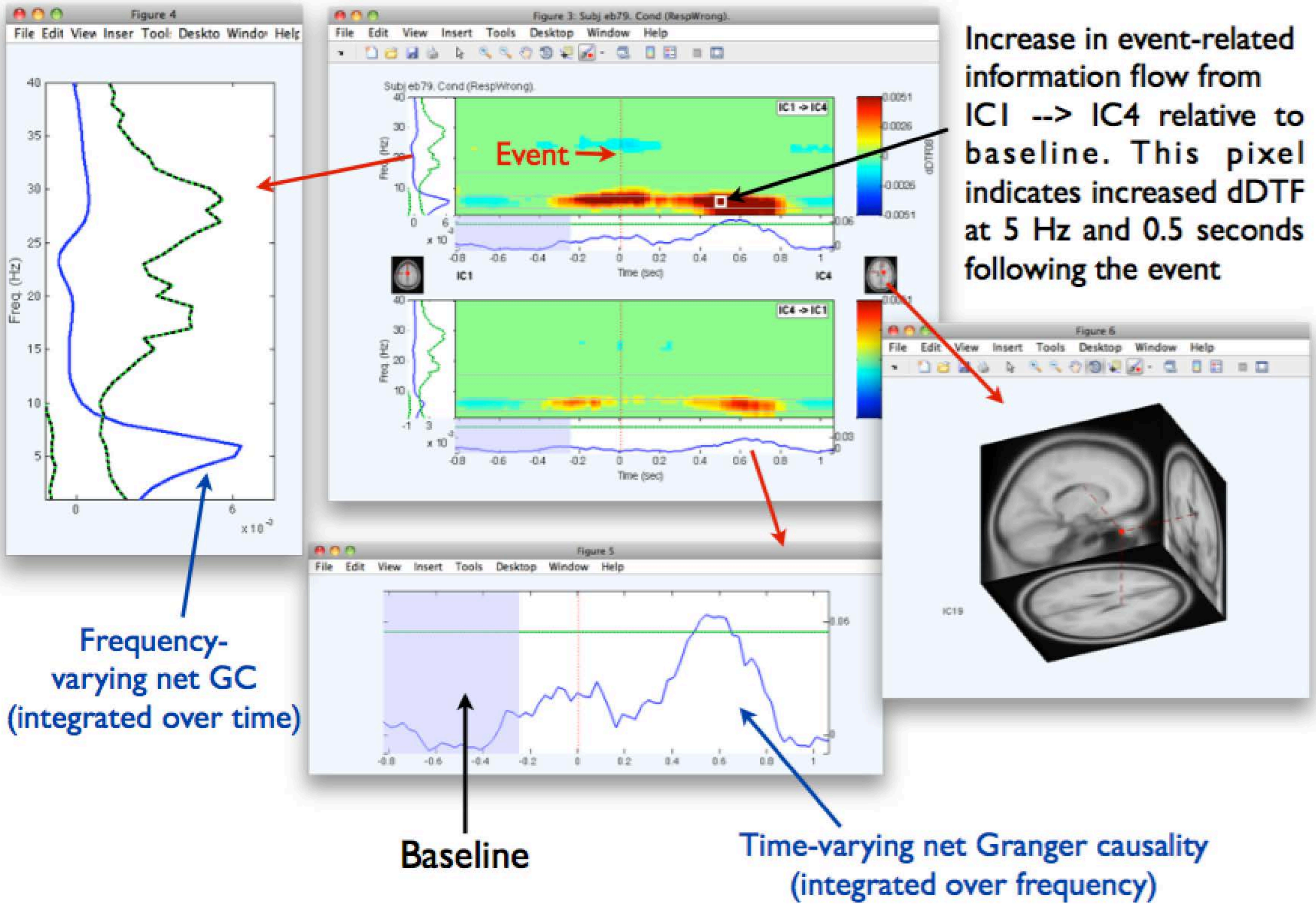
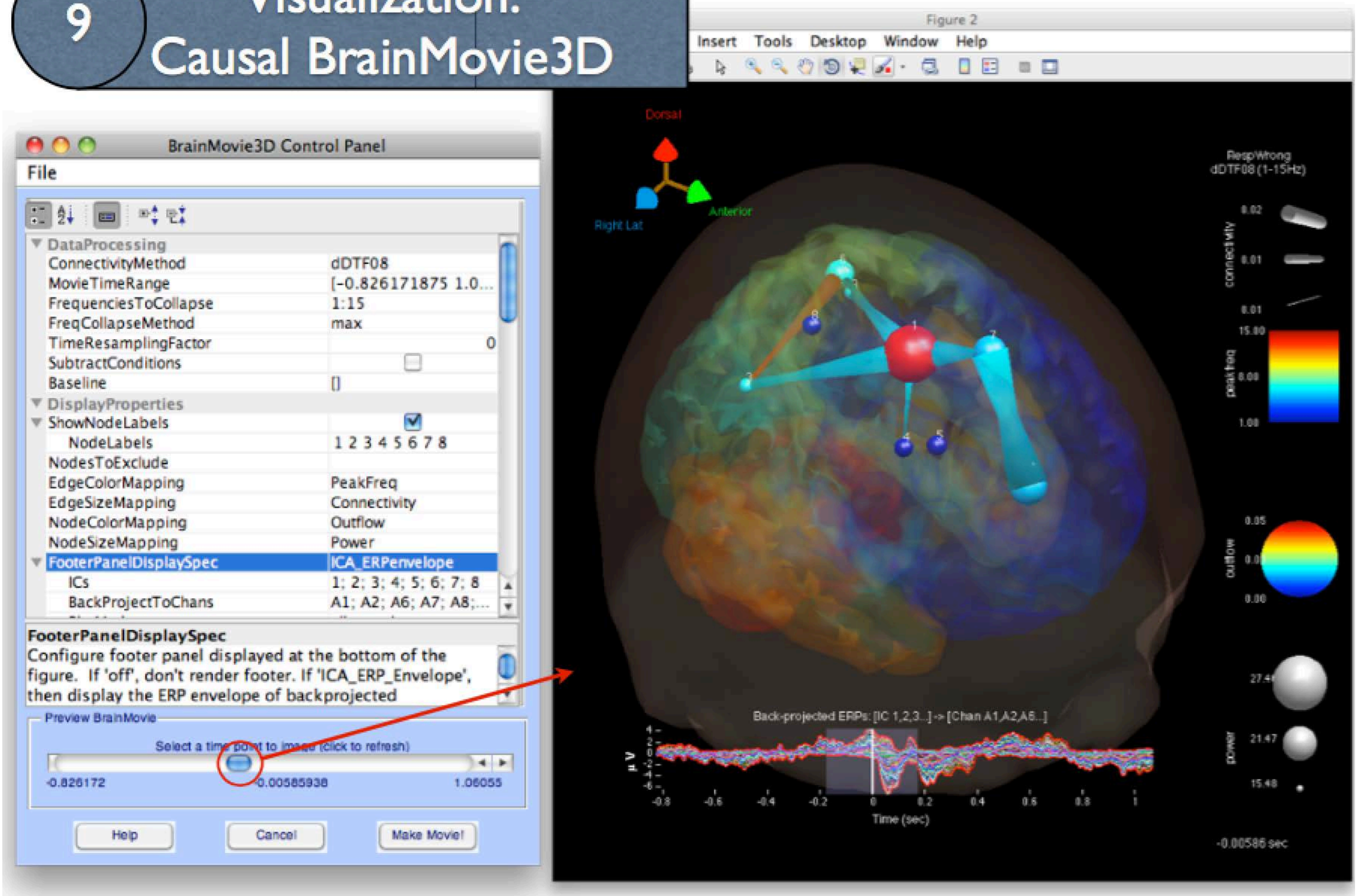



Figure 2: Subj eb79. Cond (RespWrong).
File Edit View Insert Tools Desktop Window Help
Granger Causality on off-diagonal
ERSP on diagonal
FROM
upper: dDTF08
diag: S
lower: dDTF08
1 2 3 4 5 6 7 8
TO
Frequency (Hz)
40 30 20 10
-0.5 0 0.5 1
Time (sec)



Increase in event-related information flow from IC1 --> IC4 relative to baseline. This pixel indicates increased dDTF at 5 Hz and 0.5 seconds following the event
Event
Figure 4
Figure 3: Subj eb79. Cond (RespWrong).
Subj eb79. Cond (RespWrong).
IC1 -> IC4
IC4 -> IC1
IC1
IC4
Time (sec)
Freq. (Hz)
Figure 5
Figure 6
IC19
Frequency-varying net GC (integrated over time)
Baseline
Time-varying net Granger causality (integrated over frequency)

9 Visualization: Causal BrainMovie3D



History of group-level SIFT

- Approaches
  - Tim Mullen & Wes Thompson (since 2010) '**Hierarchical Bayesian Modeling**' that interpolate missing values (i.e. inconsistency in dipole locations across subjects).
- ROI-based approaches
  - Iversen, et al, 2014: project IC activation onto cortical surface and define activity in anatomically defined cortical ROIs.
  - Nima Bigdely-Shamlo (in his PhD dissertation in 2014) **'Network Projection'** that uses dipole density and anatomical ROI. (Makoto's talk)